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 expected running time of $O(n \log(n))$.”
- “The vertex has an expected degree of $c$.”
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?

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?

![Graph showing probability distribution for hairs in soup: 0, 1, or 2 hairs equally likely]
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

Probability

Hairs in Soup

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?
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?
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 Diagram]

Every soup contains 1 hair

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

Some soups 10 hairs, most have none

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?

“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.”

Knowing that the expected value is 1 hair:

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

Some soups 10 hairs, most have none
**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."

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?  
  - Not at all  
  - Somewhat
- How likely is it that I get less than 2?  
  - Extremely
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 \textit{expected} running time of $O(n \log(n))$."
“The vertex has an \textit{expected} degree of $c$."
“In expectation there is one hair in my soup.”

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

Graph showing probability and expectation of hairs in 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?
- 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
- How likely is it that I get less than 2? Extremely

Hairs in Soup

Expectation

Probability

0 1 2 3 4 5 6 7 8 9 10

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?
- 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 exepction
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, $\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$. 
Markov’s Inequality

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

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**

![Diagram](image)
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, $\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$.

**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
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**

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

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

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

Visual Proof

\[ \mathbb{E}[X] = \sum_x x \cdot \Pr[X = x] \quad \text{fits into} \quad 3 \cdot \Pr[X \geq 3] \]
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

\[ 4 \cdot \Pr[X \geq 4] \]
Markov’s Inequality

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

**Visual Proof**

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

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

### Visual Proof

![Visual Proof Diagram](image)

**Proof**

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

fits into

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**

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

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

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

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

**Proof**

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

**Visual Proof**

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

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

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

**Visual Proof**

**Proof**

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

\[
E[X] = \sum_x x \cdot \Pr[X = x] \geq a \cdot \Pr[X \geq a]
\]

**Proof**

\[
E[X] = E[X | X < a] \cdot \Pr[X < a] + E[X | X \geq a] \cdot \Pr[X \geq a] \geq a \cdot \Pr[X \geq a] \checkmark
\]

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

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

**Proof**

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

**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?
- How likely is it 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, $\Pr[X \geq a] \leq \mathbb{E}[X]/a$.

**Visual Proof**

**Proof**

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

fits into

**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 it 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, $\Pr[X \geq a] \leq \mathbb{E}[X]/a$.

**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? $\Pr[X < 2] = 1 - \Pr[X \geq 2] \geq 1 - 1/2 = 1/2$

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: \( X \sim \text{Bin}(20, \frac{1}{5}) \)
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?
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?

\[X = 8\]

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

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

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

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

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

\[
20 \cdot \frac{1}{5} = 4
\]
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?
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?
  \[
  \Pr[X \geq 16] = \sum_{k=16}^{20} \binom{20}{k} \left(\frac{1}{5}\right)^k \cdot \left(1 - \frac{1}{5}\right)^{20-k} \approx 0.000000138
  \]
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?

\[
\Pr[X \geq 16] = \sum_{k=16}^{20} \binom{20}{k} \left( \frac{1}{5} \right)^k \cdot \left( 1 - \frac{1}{5} \right)^{20-k} \approx 0.000000138
\]
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?
  \[
  \Pr[X \geq 16] \leq \frac{\mathbb{E}[X]}{16} = 0.25
  \]

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

  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?

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

- 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?

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

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: \( \Pr[Y \geq 1] = \Pr[Y = 1] = \frac{1}{2} \)
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?
  - 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?

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

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

- 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: \(\Pr[Y \geq 1] = \Pr[Y = 1] = \frac{1}{2}\)
- Markov: \(\Pr[Y \geq 1] \leq \mathbb{E}[Y]/1 = \mathbb{E}[Y] = \frac{1}{2}\)

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?

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

Fair Coin

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

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

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

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

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?

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

Fair Coin

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

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

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

\[
\Rightarrow \text{There is no better bound (that relies only on the expected value)}
\]
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\}

\[
\Pr[D = k]
\]
\[
\Pr[U = k]
\]
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 $\{-5, -4, \ldots, 5\}$

\[
\text{Pr}[D = k]
\]

more concentr.

less concentr.

\[
\text{Pr}[U = k]
\]

Probability

$D$

$U$

$X; Y$ 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
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?

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

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?

$\mathbb{E}[D] = \sum_k \Pr[D = k] \cdot k = 0$
$\mathbb{E}[U] = \sum_k \Pr[U = k] \cdot k = 0$
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)
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?**

\[
\mathbb{E}[D] = \sum_k \Pr[D = k] \cdot k = 0 \quad \text{(same value, different shapes)}
\]

\[
\mathbb{E}[U] = \sum_k \Pr[U = k] \cdot k = 0 \quad \text{(also just seen with Markov: $\mathbb{E}$ not enough)}
\]

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

![Probability distribution graph]
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
⇒ 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. 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

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

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

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?**

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

⇒ Fix: absolute value

\[
\mathbb{E}[|D|] = \sum_k \Pr[D = k] \cdot |k| \approx 1.945
\]
Smaller expected distance to \(\mathbb{E}\)

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

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

**Problem:** + & − terms cancel

\(\Rightarrow\) Fix: absolute value

\[
\mathbb{E}[|D|] = \sum_k \Pr[D = k] \cdot |k| \approx 1.945 \quad \text{More concentrated!}
\]
\[
\mathbb{E}[|U|] = \sum_k \Pr[U = k] \cdot |k| \approx 2.727 \quad \text{Smaller expected distance to } \mathbb{E}
\]
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 $\{-5, -4, \ldots, 5\}$
- 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

⇒ Fix: absolute value

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

More concentrated!

Problem: Nobody likes absolute value
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 \([-5, -4, \ldots, 5]\)

**Expectation?**
- \(\mathbb{E}[D] = \sum_k \Pr[D = k] \cdot k = 0\) 
- \(\mathbb{E}[U] = \sum_k \Pr[U = k] \cdot k = 0\)

**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

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 \( \{-5, -4, \ldots, 5\} \)
- 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

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

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

More concentrated!

⇒ Fix: square instead

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

\[
\mathbb{E}[U^2] = \sum_k \Pr[U = k] \cdot k^2 = 10.0
\]
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?**

- $f(k) = k$
- $\mathbb{E}[D] = \sum_k \Pr[D = k] \cdot k = 0$
- $\mathbb{E}[U] = \sum_k \Pr[U = k] \cdot k = 0$ (also just seen with Markov: $\mathbb{E}$ not enough)

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

- Fix: absolute value $f(k) = |k|$
- $\mathbb{E}[|D|] = \sum_k \Pr[D = k] \cdot |k| \approx 1.945$
- $\mathbb{E}[|U|] = \sum_k \Pr[U = k] \cdot |k| \approx 2.727$

- Distance to $\mathbb{E}$

**Problem:** Nobody likes absolute value

- Fix: square instead $f(k) = k^2$
- $\mathbb{E}[D^2] = \sum_k \Pr[D = k] \cdot k^2 \approx 5.833$
- $\mathbb{E}[U^2] = \sum_k \Pr[U = k] \cdot k^2 = 10.0$

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

**These are just expectations of functions of random variables!**
Do you have a Moment?

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

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

These turn out to be particularly useful!

Moments

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

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

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. $f(X) = X^1$, $f(X) = |X|$, $f(X) = X^2$, $f(X) = \sqrt{X}$, $f(X) = X^3$, $f(X) = e^X$
- $\mathbb{E}[f(X)] = \sum_{x \in S} \Pr[X = x] \cdot f(x)$

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 $\mathbb{E}[X]$

What if $\mathbb{E}[X] \neq 0$?

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

- 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
Do you have a Moment?

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

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

$$\mathbb{E}[(X - \mathbb{E}[X])^2] = \text{Var}[X]$$
Do you have a Moment?

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

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 $E[X]$ 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

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

The smaller the variance, the more concentrated the random variable
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 $E[X^n]$.

- 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

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

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

**Proof**

\[
\Pr[|X - \mathbb{E}[X]| \geq b] = \Pr[(X - \mathbb{E}[X])^2 \geq b^2] \geq 0
\]

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

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

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

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

**Proof**

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

**Application: Unfair Coins**

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

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

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

$$\begin{align*}
\Pr[|X - \mathbb{E}[X]| \geq b] &= \Pr\left((X - \mathbb{E}[X])^2 \geq b^2\right) \\
&\leq \mathbb{E}\left[(X - \mathbb{E}[X])^2\right]/b^2 = \frac{\text{Var}[X]}{b^2} \quad \checkmark
\end{align*}$$

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

Application: Unfair Coins

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

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

- Chebychev:
  - $\Pr[X \geq 16]$
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**

$$\Pr[|X - \mathbb{E}[X]| \geq b] = \Pr[(X - \mathbb{E}[X])^2 \geq b^2] \leq \mathbb{E}[(X - \mathbb{E}[X])^2]/b^2 = \text{Var}[X]/b^2 \geq 0$$

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

**Application: Unfair Coins**

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

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

Application: Unfair Coins

- $X \sim \text{Bin}(20, \frac{1}{5})$, $\Pr[X \geq 16] \approx 0.000000138$
- $\Pr[X \geq 16] = \sum_{k=16}^{20} \binom{20}{k} \left( \frac{1}{5} \right)^k \cdot \left( 1 - \frac{1}{5} \right)^{20-k} \approx 0.000000138$
- Markov: $\Rightarrow \Pr[X \geq 16] \leq \frac{\mathbb{E}[X]}{16} = 0.25$
- Chebychev: $\Pr[X \geq 16] \leq \frac{\text{Var}[X]}{b^2}$

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

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

$$\mathbb{E}[X] = 20 \cdot \frac{1}{5} = 4$$

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

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

- Chebychev:

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

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

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

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

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

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

- Chebychev:

  $\Pr[X \geq 16]$

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

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

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

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

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

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

- Chebychev:

$$\Pr[X \geq 16] \leq \Pr[X \geq 16 \lor X \leq -8]$$

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

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

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

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

- Chebychev:

  $$\Pr[X \geq 16] \leq \Pr[X \geq 16 \lor X \leq -8] = \Pr[|X - \mathbb{E}[X]| \geq 12]$$

  $$\iff X \geq 16 \land X \leq -8$$
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**

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

**Application: Unfair Coins**

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

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

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

- Chebychev:

  $\Pr[X \geq 16] \leq \Pr[X \geq 16 \lor X \leq -8]$

  $= \Pr[|X - \mathbb{E}[X]| \geq 12]$

  $\leq \frac{\text{Var}[X]}{12^2}$

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

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

Application: Unfair Coins

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

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

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

- Chebychev:

  $\Pr[X \geq 16] \leq \Pr[X \geq 16 \vee X \leq -8]$

  $= \Pr[|X - \mathbb{E}[X]| \geq 12]$

  $\leq \frac{\text{Var}[X]}{12^2}$

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

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

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

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

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

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

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

$X \geq 12$ 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

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

Application: Unfair Coins

- $X \sim \text{Bin}(20, \frac{1}{5})$, $\Pr[X \geq 16] \approx 0.000000138$
- Markov: $\Rightarrow \Pr[X \geq 16] \leq \frac{\mathbb{E}[X]}{16} = 0.25$.
- Chebychev:
  $$\Pr[X \geq 16] \leq \Pr[X \geq 16 \vee X \leq -8] = \Pr[|X - \mathbb{E}[X]| \geq 12] \leq \frac{\text{Var}[X]}{12^2}$$

$$X \sim \text{Bin}(n, p) : \text{Var}[X] = np(1 - p)$$
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, $\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$.

Proof

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

Application: Unfair Coins

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

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

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

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

$X \geq 16 \text{ or } X \leq -8$
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**

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

**Application: Unfair Coins**

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

- Markov: $\Rightarrow \Pr[X \geq 16] \leq \frac{\mathbb{E}[X]}{16} = 0.25$
- Chebychev:
  \[
  \Pr[X \geq 16] \leq \Pr[X \geq 16 \lor X \leq -8]
  = \Pr[|X - \mathbb{E}[X]| \geq 12]
  \leq \frac{\text{Var}[X]}{12^2} = \frac{16}{5\cdot 144}
  \]

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

- $X \geq 16 \iff X - \mathbb{E}[X] \geq 16 - \mathbb{E}[X]$\[
  \iff X - \mathbb{E}[X] \geq 12 \implies X \geq 16 \text{ or } X \leq -8
  \]
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**

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

**Application: Unfair Coins**

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

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

- Chebychev:
  $$
  \Pr[X \geq 16] \leq \Pr[X \geq 16 \lor X \leq -8] = \Pr[|X - \mathbb{E}[X]| \geq 12] \leq \frac{\text{Var}[X]}{12^2} = \frac{16}{5 \cdot 144} \approx 0.022
  $$
  **Order of magnitude better than Markov!**

$X \sim \text{Bin}(n, p)$ : $\text{Var}[X] = np(1 - p)$
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 $v$: $\deg(v) \sim \text{Bin}(n - 1, p)$
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 $v$: $\text{deg}(v) \sim \text{Bin}(n-1, p)$
- 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 $d_{TV}(\text{deg}(v), X) = o(1)$
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 \( v \): \( \text{deg}(v) \sim \text{Bin}(n - 1, p) \)
- 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 \( G = (V, E) \)
  \[
  N_d = \sum_{v \in V} \mathbb{1}_{\{\text{deg}(v) = d\}} 
  \]
  (normalized: \( \frac{1}{n} N_d \), for \( n = |V| \))

\[\begin{align*}
\text{Pr}[X = d], & \quad n = 100 \\
\text{Pr}[X = d], & \quad n = 1000 \\
\text{Pr}[X = d], & \quad n = 10000
\end{align*}\]
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

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

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

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

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

- **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 $\epsilon > 0$ 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 $\frac{1}{n}N_d$

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

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

- **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 $\frac{1}{n} N_{d}$
\[
\lim_{n \to \infty} \left| \Pr[X = d] - \mathbb{E} \left[ \frac{1}{n} N_{d} \right] \right| = 0
\]
\[
= \frac{1}{n} \mathbb{E}[N_{d}]
\]
\[
= \frac{1}{n} \mathbb{E} \left[ \sum_{v \in V} 1_{\{\deg(v) = d\}} \right]
\]
\[
= \frac{1}{n} \sum_{v \in V} \mathbb{E}[1_{\{\deg(v) = d\}}]
\]

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

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

- **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 $\frac{1}{n} N_d$

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

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

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

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

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

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

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

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

  $\lambda = c + O(1/n) \to c$ 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 $\epsilon > 0$ 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 $\frac{1}{n} N_d$

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

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

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

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

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

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

- **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 $\frac{1}{n} N_d$

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

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

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

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

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

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

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

- **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

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

**Proof**

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

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

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

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

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

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

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

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

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

  $d_{TV}(X, \deg(v)) = o(1)$

  (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 $\frac{1}{n} N_d$  
  \[
  \lim_{n \to \infty} \left| \Pr[X = d] - \mathbb{E} \left[ \frac{1}{n} N_d \right] \right| = 0 \checkmark
  \]
  
  \[
  \left| \Pr[X = d] - \mathbb{E} \left[ \frac{1}{n} N_d \right] \right| = \left| \Pr[X = d] - \Pr[\deg(v) = d] \right| \leq \sum_{d \geq 0} \left| \Pr[X = d] - \Pr[\deg(v) = d] \right|
  \]
  
  \[
  = 2 \cdot d_{TV}(X, \deg(v)) = o(1) \quad n \to \infty \to 0 \checkmark
  \]

- **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 $\frac{1}{n}N_d$

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

$\textbf{Chebychev:}$ $X$ 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 $\frac{1}{n} N_d$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

**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 $\frac{1}{n}N_d$

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

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

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

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

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

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

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

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 $\frac{1}{n} N_d$

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

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

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

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

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

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

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

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

Lin. of Exp.

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

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

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

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

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

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

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

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

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

Indicator RV $X$: $X^2 = X$,

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

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

$= 1$ if $\deg(v) = d \land \deg(u) = d$

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 $\frac{1}{n} N_d$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

\[= \text{Chebychev: } X \text{ finite variance, } b > 0\]

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

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

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

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

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

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

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

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

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

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

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

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

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

**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 $\frac{1}{n}N_d$

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

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

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

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

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

$$\leq 1$$

**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 $\frac{1}{n} N_d$

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

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

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

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

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

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

$$- \Pr[\deg(v) = d]^2 \leq 1$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

**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 $\frac{1}{n}N_d$

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

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

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

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

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

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

**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 $\frac{1}{n} N_d$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

\[ \deg(v) \overset{d}{=} \deg(u) \]

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

**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 $\frac{1}{n} N_d$

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

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

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

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

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

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

**Couplings**

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

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

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

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

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

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

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

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

**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 $\text{deg}(v)$
- $Y_1, Y_2 \sim \text{Bin}(n-2, p)$
- $X_1, X_2 \sim \text{Ber}(p)$

\begin{itemize}
  \item $u$
  \item $v$
\end{itemize}
Step 2: Concentration of $\frac{1}{n}N_d$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

**Couplings**

- Consider $\text{deg}(u)$ and $\text{deg}(v)$
- $Y_1, Y_2 \sim \text{Bin}(n - 2, p)$ independent
- $X_1, X_2 \sim \text{Ber}(p)$
Step 2: Concentration of \( \frac{1}{n} N_d \)

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

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

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

**Couplings**

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

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

**Chebychev:** \( X \) finite variance, \( b > 0 \)
\[ \text{Pr} \left[ |X - \mathbb{E}[X]| \geq b \right] \leq \frac{\text{Var}[X]}{b^2} \]

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

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

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

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

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

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

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

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

Couplings

- Consider $\deg(u)$ and $\deg(v)$
- $Y_1, Y_2 \sim \text{Bin}(n - 2, p)$ independent
- $X_1, X_2 \sim \text{Ber}(p)$
Step 2: Concentration of $\frac{1}{n} N_d$

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

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

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

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

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

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

- $Y_1, Y_2 \sim \text{Bin}(n-2, p)$
- $X_1, X_2 \sim \text{Ber}(p)$

**Couplings**

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

---

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

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

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

**Consider**

- $\text{deg}(u)$ and $\text{deg}(v)$
- $Y_1, Y_2 \sim \text{Bin}(n-2, p)$
- $X_1, X_2 \sim \text{Ber}(p)$
Step 2: Concentration of $\frac{1}{n} N_d$

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

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

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

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

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

**Couplings**

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

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

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

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

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

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

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

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

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

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

- $Y_1, Y_2 \sim \text{Bin}(n - 2, p)$
- $X_1, X_2 \sim \text{Ber}(p)$
- $(\text{deg}(v), \text{deg}(u)) \equiv (X_1 + Y_1, X_1 + Y_2)$

**Couplings**

$\sim$ independent

$\equiv$ dependent
Step 2: Concentration of $\frac{1}{n} N_d$

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

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

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

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

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

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

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

**Couplings**
- Consider $\text{deg}(u)$ and $\text{deg}(v)$
- $Y_1, Y_2 \sim \text{Bin}(n-2, p)$
- $X_1, X_2 \sim \text{Ber}(p)$
- $(\text{deg}(v), \text{deg}(u)) \overset{d}{=} (X_1+Y_1, X_1+Y_2)$

\[
\begin{align*}
\text{deg}(v) &\overset{\text{dependent}}{\parallel} \text{deg}(u) \\
X_1+Y_1 &\overset{\text{dependent}}{\parallel} X_1+Y_2
\end{align*}
\]
Step 2: Concentration of $\frac{1}{n} N_d$

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

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

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

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

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

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

\begin{align*}
\text{Chebychev: } &X \text{ finite variance, } b > 0 \\
&\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}
\end{align*}

\begin{align*}
(\sum_i a_i)^2 &= \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j
\end{align*}

**Couplings**

- Consider $\text{deg}(u)$ and $\text{deg}(v)$
- $Y_1, Y_2 \sim \text{Bin}(n-2, p)$ independent
- $X_1, X_2 \sim \text{Ber}(p)$
- $(\text{deg}(v), \text{deg}(u)) \overset{d}{=} (X_1+Y_1, X_1+Y_2)$
- $Y_1$ $\overset{\alpha}{\leftarrow} X_1 + Y_1$
- $\overset{\alpha}{\leftarrow} X_1 + Y_1$
- $\overset{\alpha}{\leftarrow} X_2 + Y_2$
- $\overset{\alpha}{\leftarrow} X_2 + Y_2$
Step 2: Concentration of $\frac{1}{n}N_d$

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

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

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

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

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

**Couplings**

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

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

**Chebychev:** $X$ finite variance, $b > 0$
\[ \Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2} \]

\[ (\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j \]
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 $\text{deg}(u)$ and $\text{deg}(v)$
- $Y_1, Y_2 \sim \text{Bin}(n-2, p)$
- $X_1, X_2 \sim \text{Ber}(p)$
- $(\text{deg}(v), \text{deg}(u)) \overset{d}{=} (X_1 + Y_1, X_1 + Y_2)$

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

**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 $\frac{1}{n}N_d$

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

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

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

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

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

**Couplings**
- Consider $\text{deg}(u)$ and $\text{deg}(v)$
- $Y_1, Y_2 \sim \text{Bin}(n - 2, p)$
- $X_1, X_2 \sim \text{Ber}(p)$
- $(\text{deg}(v), \text{deg}(u)) \overset{d}{=} (X_1 + Y_1, X_1 + Y_2)$

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

**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 $\frac{1}{n}N_d$

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

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

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

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

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

$$= \frac{1}{n} + \text{Pr}[X_1 + Y_1 = d \land X_1 + Y_2 = d]$$

**Couplings**

- Consider $\text{deg}(u)$ and $\text{deg}(v)$.
- $Y_1, Y_2 \sim \text{Bin}(n - 2, p)$
- $X_1, X_2 \sim \text{Ber}(p)$
- $(\text{deg}(v), \text{deg}(u)) \overset{d}{=} (X_1 + Y_1, X_1 + Y_2)$

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

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

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

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

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

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

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

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

$$= \frac{1}{n} + \Pr \left[ X_1 + Y_1 = d \land X_1 + Y_2 = d \right]$$

$$- \Pr \left[ X_1 + Y_1 = d \right] \Pr \left[ X_2 + Y_2 = d \right]$$

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

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

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

**Couplings**

- Consider $\text{deg}(u)$ and $\text{deg}(v)$
- $Y_1, Y_2 \sim \text{Bin}(n - 2, p)$
- $X_1, X_2 \sim \text{Ber}(p)$
- $(\text{deg}(v), \text{deg}(u)) \overset{d}{=} (X_1 + Y_1, X_1 + Y_2)$

$$(\text{deg}(v), \text{deg}(u)) \overset{d}{=} (X_1 + Y_1, X_1 + Y_2)$$
Step 2: Concentration of $\frac{1}{n} N_d$

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

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

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

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

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

$$= \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d]$$

$$- \Pr[X_1 + Y_1 = d] \Pr[X_2 + Y_2 = d]$$

**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 $\frac{1}{n}N_d$

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

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

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

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

\[
= \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d] - \Pr[X_1 + Y_1 = d] \Pr[X_2 + Y_2 = d]
\]

\[
= \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d] - \Pr[X_1 + Y_1 = d \land X_2 + Y_2 = d]
\]

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

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

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

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

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

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

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

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2 = \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right) \leq \frac{1}{n} + \Pr[\deg(v) = d \land \deg(u) = d] - \Pr[\deg(v) = d] \Pr[\deg(u) = d] = \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d] - \Pr[X_1 + Y_1 = d] \Pr[X_2 + Y_2 = d] \leq \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d] - \Pr[X_1 + Y_1 = d] \Pr[X_2 + Y_2 = d] \leq \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d] - \Pr[X_1 + Y_1 = d] \Pr[X_2 + Y_2 = d]$$

Fréchet: $\Pr[A] - \Pr[B] \leq \Pr[A \land \neg B]$

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

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

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

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

Consider $\deg(u)$ and $\deg(v)$

- $Y_1, Y_2 \sim \text{Bin}(n - 2, p)$
- $X_1, X_2 \sim \text{Ber}(p)$
- $(\deg(v), \deg(u)) \overset{d}{=} (X_1 + Y_1, X_1 + Y_2)$
- $\deg(v) \overset{\sim}{\not\parallel} \deg(u)$
- $X_1 + Y_1 \overset{\sim}{\not\parallel} X_1 + Y_2$
Step 2: Concentration of $\frac{1}{n} N_d$

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

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

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

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

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

$$= \frac{1}{n} + \text{Pr}[X_1 + Y_1 = d \land X_1 + Y_2 = d]$$

$$- \text{Pr}[X_1 + Y_1 = d] \text{Pr}[X_2 + Y_2 = d]$$

$$= \frac{1}{n} + \text{Pr}[X_1 + Y_1 = d \land X_1 + Y_2 = d]$$

$$- \text{Pr}[X_1 + Y_1 = d \land X_2 + Y_2 = d]$$

$$\leq \frac{1}{n} + \text{Pr}[X_1 + Y_1 = d \land X_1 + Y_2 = d]$$

$$\land \left( X_1 + Y_1 \neq d \lor X_2 + Y_2 \neq d \right)$$

For the whole event to occur, this needs to happen, which excludes this from happening.

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

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

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

Couplings

- Consider $\text{deg}(u)$ and $\text{deg}(v)$
- $Y_1, Y_2 \sim \text{Bin}(n - 2, p)$
- $X_1, X_2 \sim \text{Ber}(p)$
- $(\text{deg}(v), \text{deg}(u)) \overset{d}{=} (X_1 + Y_1, X_1 + Y_2)$

Fréchet: $\text{Pr}[A] - \text{Pr}[B] \leq \text{Pr}[A \land \bar{B}]$
Step 2: Concentration of $\frac{1}{n} N_d$

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

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

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

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

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

$$= \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d]$$

$$- \Pr[X_1 + Y_1 = d] \Pr[X_2 + Y_2 = d]$$

$$= \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d]$$

$$- \Pr[X_1 + Y_1 = d \land X_2 + Y_2 = d]$$

$$\leq \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land (X_1 + Y_1 \neq d \lor X_2 + Y_2 \neq d)]$$

$$= \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d]$$

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

**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:** $\Pr[A] - \Pr[B] \leq \Pr[A \land \overline{B}]$
Step 2: Concentration of $\frac{1}{n} N_d$

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

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

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

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

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

$$\leq \frac{1}{n} + \Pr \left[ X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d \right]$$

Fréchet: \( \Pr[A] - \Pr[B] \leq \Pr[A \land \bar{B}] \)

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

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

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

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

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

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

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

\[
\leq \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d]
\]

\[
= \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \Pr[X_1 = 0]
\]

\[
+ \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 1] \Pr[X_1 = 1]
\]

**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:** $\Pr[A] - \Pr[B] \leq \Pr[A \land \bar{B}]$

Law of total probability

- $Y_1, Y_2 \sim \text{Bin}(n - 2, p)$
- $X_1, X_2 \sim \text{Ber}(p)$ independent
Step 2: Concentration of \( \frac{1}{n} N_d \)

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

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

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

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

\[
\leq \frac{1}{n} + \Pr \left[ X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d \right]
\]

\[
= \frac{1}{n} + \Pr \left[ X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d \mid X_1 = 0 \right] \Pr \left[ X_1 = 0 \right] + \Pr \left[ X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d \mid X_1 = 1 \right] \Pr \left[ X_1 = 1 \right]
\]

\[
\leq \frac{1}{n} + \Pr \left[ X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d \mid X_1 = 0 \right] \Pr \left[ X_1 = 0 \right] + \Pr \left[ X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d \mid X_1 = 1 \right] \Pr \left[ X_1 = 1 \right]
\]

\[
\leq 1
\]

\[
\text{Chebychev: } X \text{ finite variance, } b > 0 \quad \Pr \left[ |X - \mathbb{E}[X]| \geq b \right] \leq \frac{\text{Var}[X]}{b^2}
\]

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

\[
\text{Fréchet: } \Pr[A] - \Pr[B] \leq \Pr[A \land \neg B]
\]

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

\[
X_1, X_2 \sim \text{Ber}(p) \quad \text{independent}
\]

\[
Y_1, Y_2 \sim \text{Bin}(n - 2, p) \quad \text{independent}
\]
Step 2: Concentration of $\frac{1}{n}N_d$

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

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

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

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

\[
\leq \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d]
\]

\[
= \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \Pr[X_1 = 0] + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 1] \Pr[X_1 = 1]
\]

\[
\leq \frac{1}{n} + \Pr[X_1 = 0] + \Pr[X_1 = 1]
\]

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

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

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

Fréchet: $\Pr[A] - \Pr[B] \leq \Pr[A \land \bar{B}]$

Law of total probability

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

$Y_1, Y_2 \sim \text{Bin}(n - 2, p)$

$X_1, X_2 \sim \text{Ber}(p)$ independent
Step 2: Concentration of $\frac{1}{n} N_d$

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

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

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

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

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

$$\leq \frac{1}{n} + \Pr [X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d]$$

$$= \frac{1}{n} + \Pr [X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \Pr [X_1 = 0]$$

$$+ \Pr [X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 1] \Pr [X_1 = 1]$$

$$\leq \frac{1}{n} + \Pr [X_1 = 1]$$

$$\leq 1$$

Law of total probability

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

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

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

$\text{Fréchet:} \ \Pr[A] - \Pr[B] \leq \Pr[A \land \bar{B}]$
Step 2: Concentration of $\frac{1}{n} N_d$

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

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

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

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

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

$$\leq \frac{1}{n} + \Pr [X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d]$$

$$= \frac{1}{n} + \Pr [X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \Pr [X_1 = 0]$$

$$+ \Pr [X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 1] \Pr [X_1 = 1]$$

$$\leq \frac{1}{n} + \Pr [Y_1 = d \land Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \leq 1$$

$$+ \Pr [X_1 = 1]$$

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

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

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

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

**Fréchet:** $\Pr[A] - \Pr[B] \leq \Pr[A \land \bar{B}]$

Law of total probability

- $Y_1, Y_2 \sim \text{Bin}(n - 2, p)$
- $X_1, X_2 \sim \text{Ber}(p)$
Step 2: Concentration of $\frac{1}{n} N_d$

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

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

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

\[
\leq \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d] \leq \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \Pr[X_1 = 0] + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 1] \Pr[X_1 = 1] \leq 1
\]

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

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

Fréchet: $\Pr[A] - \Pr[B] \leq \Pr[A \land \overline{B}]$

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

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

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

Law of total probability

\[
\begin{align*}
\Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] &\leq \Pr[Y_1 = d \land Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \leq 1 \\
\Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 1] &\leq 1 \\
\end{align*}
\]

\[
\begin{align*}
Y_1, Y_2 &\sim \text{Bin}(n - 2, p) \\
X_1, X_2 &\sim \text{Ber}(p)
\end{align*}
\]

\[
\text{independent}
\]
Step 2: Concentration of \( \frac{1}{n} N_d \)

\[
\begin{align*}
\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \epsilon \right] & \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\epsilon^2} \\
\text{Var} \left[ \frac{1}{n} N_d \right] & = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2 \\
& = \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right) \\
& \leq \frac{1}{n} + \Pr[\deg(v) = d \land \deg(u) = d] - \Pr[\deg(v) = d] \Pr[\deg(u) = d] \quad \text{deg}(v) \overset{d}{=} \text{deg}(u) \\
& \leq \frac{1}{n} + \Pr[ X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d] \\
& = \frac{1}{n} + \Pr[ X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \Pr[ X_1 = 0] \\
& \quad + \Pr[ X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 1] \Pr[ X_1 = 1] \\
& \leq \frac{1}{n} + \Pr[ Y_1 = d \land Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \\
& \quad + \Pr[ X_1 = 1] \\
& \Rightarrow X_2 = 1
\end{align*}
\]

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

**Chebychev:** \( X \) finite variance, \( b > 0 \)

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

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

**Fréchet:** \( \Pr[A] - \Pr[B] \leq \Pr[A \land \bar{B}] \)

\[
\begin{align*}
\text{Law of total probability}
\end{align*}
\]

\[
\begin{align*}
Y_1, Y_2 & \sim \text{Bin}(n-2; p) \\
X_1, X_2 & \sim \text{Ber}(p)
\end{align*}
\]

\[
\begin{align*}
\text{independent}
\end{align*}
\]
Step 2: Concentration of $\frac{1}{n}N_d$

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

\[
\text{Var} \left[ \frac{1}{n}N_d \right] = \mathbb{E} \left[ (\frac{1}{n}N_d)^2 \right] - \mathbb{E} \left[ \frac{1}{n}N_d \right]^2
\]
\[
= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right)
\]
\[
\leq \frac{1}{n} + \Pr[\text{deg}(v) = d \land \text{deg}(u) = d] - \Pr[\text{deg}(v) = d] \Pr[\text{deg}(u) = d] \quad \text{deg}(v) \overset{d}{=} \text{deg}(u)
\]
\[
\leq \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d]
\]
\[
= \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \Pr[X_1 = 0]
\]
\[
+ \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 1] \Pr[X_1 = 1]
\]
\[
\leq \frac{1}{n} + \Pr[Y_1 = d \land Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \Rightarrow X_2 = 1
\]
\[
+ \Pr[X_1 = 1]
\]
\[
= \frac{1}{n} + \Pr[Y_1 = d \land Y_2 = d \land X_2 = 1 | X_1 = 0] + \Pr[X_1 = 1]
\]
Step 2: Concentration of $\frac{1}{n} N_d$

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

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

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

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

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

$$\leq \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d]$$

$$= \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \Pr[X_1 = 0]$$

$$+ \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 1] \Pr[X_1 = 1]$$

$$\leq \frac{1}{n} + \Pr[Y_1 = d \land Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \leq 1$$

$$+ \Pr[X_1 = 1] \Rightarrow X_2 = 1$$

$$= \frac{1}{n} + \Pr[Y_1 = d \land Y_2 = d \land X_2 = 1 | X_1 = 0] + \Pr[X_1 = 1]$$

$$\leq 1$$

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

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

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

Fréchet: $\Pr[A] - \Pr[B] \leq \Pr[A \land \bar{B}]$

Law of total probability

$Y_1, Y_2 \sim \text{Bin}(n - 2, p)$ independent

$X_1, X_2 \sim \text{Ber}(p)$ independent
Step 2: Concentration of $\frac{1}{n} N_d$

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

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

$$\leq \frac{1}{n} + \Pr[X_1 + Y_1 = d \wedge X_1 + Y_2 = d \wedge X_2 + Y_2 \neq d] \leq \frac{1}{n} + \Pr[Y_1 = d \wedge Y_2 = d \wedge X_2 + Y_2 \neq d | X_1 = 0] + \Pr[X_1 = 1]$$

$$\leq \frac{1}{n} + \Pr[Y_1 = d \wedge Y_2 = d \wedge X_2 + Y_2 \neq d | X_1 = 0] \leq 1$$

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: $\Pr[A] - \Pr[B] \leq \Pr[A \wedge \bar{B}]$

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 $\frac{1}{n} N_d$

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

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

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

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

$$\leq \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d]$$

$$= \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d|X_1 = 0] \Pr[X_1 = 0]$$

$$+ \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d|X_1 = 1] \Pr[X_1 = 1]$$

$$\leq \frac{1}{n} + \Pr[Y_1 = d \land Y_2 = d \land X_2 + Y_2 \neq d|X_1 = 0] \Pr[X_1 = 0]$$

$$+ \Pr[X_1 = 1]$$

$$= \frac{1}{n} + \Pr[Y_1 = d \land Y_2 = d \land X_2 = 1|X_1 = 0] \Pr[X_1 = 0]$$

$$+ \Pr[X_1 = 1]$$

$$\leq \frac{1}{n} + \Pr[X_2 = 1] + \Pr[X_1 = 1]$$

$$\leq \frac{1}{n} + \Pr[X_2 = 1] + \Pr[X_1 = 1]$$

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

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

$$\text{Fréchet: } \Pr[A] - \Pr[B] \leq \Pr[A \land \overline{B}]$$

$$\text{Law of total probability}$$
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}$

Var$[\frac{1}{n} N_d] = \mathbb{E}[(\frac{1}{n} N_d)^2] - \mathbb{E}[\frac{1}{n} N_d]^2 \leq \frac{1}{n} + 2p$

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

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

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

$\leq \frac{1}{n} + \text{Pr}[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d]$}

$= \frac{1}{n} + \text{Pr}[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \cdot \text{Pr}[X_1 = 0]$

$+ \text{Pr}[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 1] \cdot \text{Pr}[X_1 = 1]$

$\leq \frac{1}{n} + \text{Pr}[Y_1 = d \land Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \leq 1$

$+ \text{Pr}[X_1 = 1] \Rightarrow X_2 = 1$

$\leq \frac{1}{n} + \text{Pr}[X_1 = 1] \leq \frac{1}{n} + \text{Pr}[X_2 = 1] + \text{Pr}[X_1 = 1]$

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

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

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

\[
\text{Fréchet: } \text{Pr}[A] - \text{Pr}[B] \leq \text{Pr}[A \land \bar{B}]
\]

\[
\text{Law of total probability}
\]
Step 2: Concentration of $\frac{1}{n}N_d$

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

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

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

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

\[
\leq \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d]
\]

\[
= \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \Pr[X_1 = 0]
+ \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 1] \Pr[X_1 = 1]
\]

\[
\leq \frac{1}{n} + \Pr[Y_1 = d \land Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \leq 1
\]

\[
\Rightarrow X_2 = 1
\]

\[
= \frac{1}{n} + \Pr[Y_1 = d \land Y_2 = d \land X_2 = 1 | X_1 = 0] + \Pr[X_1 = 1] \leq \frac{1}{n} + \Pr[X_2 = 1] + \Pr[X_1 = 1]
\]

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

Chebychev: $X$ finite variance, $b > 0$
\[
\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}
\]

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

Fréchet: $\Pr[A] - \Pr[B] \leq \Pr[A \land \bar{B}]$

Law of total probability

\[
\begin{align*}
Y_1, Y_2 &\sim \text{Bin}(n - 2, p) \\
X_1, X_2 &\sim \text{Ber}(p)
\end{align*}
\]
Step 2: Concentration of $\frac{1}{n} N_d$

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

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2 \leq \frac{1}{n} + 2p = \frac{1}{n} + 2\frac{c}{n} \xrightarrow{n \to \infty} 0$$

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

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

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

$$\leq \frac{1}{n} + \text{Pr}[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d]$$

$$\leq \frac{1}{n} + \text{Pr}[Y_1 = d \land Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \text{Pr}[X_1 = 0]$$

$$+ \text{Pr}[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 1] \text{Pr}[X_1 = 1]$$

$$\leq \frac{1}{n} + \text{Pr}[Y_1 = d \land Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \Rightarrow X_2 = 1$$

$$\leq 1$$

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

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

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

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

**Fréchet:** $\Pr[A] - \Pr[B] \leq \Pr[A \land \bar{B}]$

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 \( \frac{1}{n} N_d \)

\[
\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2} \xrightarrow{n \to \infty} 0
\]

\[
\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2 \leq \frac{1}{n} + 2p = \frac{1}{n} + 2 \frac{c}{n} \xrightarrow{n \to \infty} 0
\]

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: \( \Pr[A] - \Pr[B] \leq \Pr[A \land \overline{B}] \)

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

Law of total probability

\[
\begin{align*}
\Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \Pr[X_1 = 0] \\
+ \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 1] \Pr[X_1 = 1] \\
\leq \frac{1}{n} + \Pr[Y_1 = d \land Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \\
+ \Pr[X_1 = 1] \\
\Rightarrow X_2 = 1
\end{align*}
\]

\[
\begin{align*}
\Pr[Y_1 = d \land Y_2 = d \land X_2 = 1 | X_1 = 0] + \Pr[X_1 = 1] \\
\leq \frac{1}{n} + \Pr[X_2 = 1] + \Pr[X_1 = 1]
\end{align*}
\]

\( Y_1, Y_2 \sim \text{Bin}(n - 2, p) \)
\( X_1, X_2 \sim \text{Ber}(p) \)

\( \text{independent} \)
Step 2: Concentration of $\frac{1}{n}N_d$

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

Var$[\frac{1}{n}N_d] = \mathbb{E}[(\frac{1}{n}N_d)^2] - \mathbb{E}[\frac{1}{n}N_d]^2 \leq \frac{1}{n} + 2p = \frac{1}{n} + 2\varepsilon \xrightarrow{n \to \infty} 0$

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

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

Fréchet: $\Pr[A] - \Pr[B] \leq \Pr[A \land \bar{B}]$

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

Fréchet: $\Pr[A] - \Pr[B] \leq \Pr[A \land \bar{B}]$

Law of total probability

$$\Pr[Y_1 = d \land Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \leq 1$$

$Y_1, Y_2 \sim \text{Bin}(n - 2, p)$

$X_1, X_2 \sim \text{Ber}(p)$

$\frac{1}{n} + \Pr[X_1 = 1] \leq \frac{1}{n} + \Pr[X_2 = 1] + \Pr[X_1 = 1]$ independent

$\Rightarrow X_2 = 1$ independent
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 $\frac{1}{n} N_d$.
  \[
  \lim_{n \to \infty} \left| \Pr[X = d] - \mathbb{E} \left[ \frac{1}{n} N_d \right] \right| = 0 \checkmark
  \]

- **Step 2:** $\frac{1}{n} N_d$ is concentrated (via Chebychev).
  \[
  \lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0 \checkmark
  \]
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 \( a > 0 \)
\[
\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}
\]
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 \( a > 0 \)
  \[
  \Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}
  \]
- 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)

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 $E[X^n]$
- We can capture *all* moments of $X$ using a single function

**Definition**: For a random variable $X$ the **moment generating function** is $M_X(t) = E[e^{tX}]$.

Looks scary, but is again just $E[f(X)]$ for $f(X) = e^{tX}$. 

---

Maximilian Katzmann, Stefan Walzer – Probability & Computing
Institute of Theoretical Informatics, Algorithm Engineering & Scalable Algorithms
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** $M_{X+Y}(t) = \mathbb{E}[e^{t(X+Y)}] = \mathbb{E}[e^{tX} \cdot e^{tY}]$

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 $M^{(n)}_X(t)$ we have $M^{(n)}_X(0) = \mathbb{E}[X^n]$ (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** $M_{X+Y}(t) = \mathbb{E}[e^{t(X+Y)}] = \mathbb{E}[e^{tX} \cdot e^{tY}] = \mathbb{E}[e^{tX}] \cdot \mathbb{E}[e^{tY}]$
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**

\[
M_{X+Y}(t) = \mathbb{E}[e^{t(X+Y)}] = \mathbb{E}[e^{tX} \cdot e^{tY}] = \mathbb{E}[e^{tX}] \cdot \mathbb{E}[e^{tY}] = M_X(t) \cdot M_Y(t)
\]
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, $\Pr[X \geq a] \leq \min_{t>0} \frac{\mathbb{E}[e^{tX}]}{e^{ta}}$ and $\Pr[X \leq a] \leq \min_{t<0} \frac{\mathbb{E}[e^{tX}]}{e^{ta}}$.

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. "A conversation with Herman Chernoff", John Bather, Statist. Sci. 1996
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, $\Pr[X \geq a] \leq \min_{t>0} \frac{\mathbb{E}[e^{tX}]}{e^{ta}}$ 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 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 \( M^{(n)}_X(t) \) we have \( M^{(n)}_X(0) = \mathbb{E}[X^n] \)

**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, \( \Pr[X \geq a] \leq \min_{t>0} \frac{\mathbb{E}[e^{tX}]}{e^{ta}} \) and \( \Pr[X \leq a] \leq \min_{t<0} \frac{\mathbb{E}[e^{tX}]}{e^{ta}}. \)

**Proof** for all \( t > 0 \):
\[
\Pr[X \geq a] = \Pr[e^{tX} \geq e^{ta}] \leq \frac{\mathbb{E}[e^{tX}]}{e^{ta}}
\]

Markov: \( X \) non-negative, \( b > 0 \):
\[
\Pr[X \geq b] \leq \frac{\mathbb{E}[X]}{b}.
\]
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 \frac{\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. \n
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 $X, Y$: $M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$.

**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, $\Pr[X \geq a] \leq \min_{t>0} \frac{\mathbb{E}[e^{tX}]}{e^{ta}}$ and $\Pr[X \leq a] \leq \min_{t<0} \frac{\mathbb{E}[e^{tX}]}{e^{ta}}$.

**Proof**

for all $t > 0$: $\Pr[X \geq a] = \Pr[e^{tX} \geq e^{ta}] \leq \frac{\mathbb{E}[e^{tX}]}{e^{ta}} \leq \min_{t>0} \frac{\mathbb{E}[e^{tX}]}{e^{ta}}$

for all $t < 0$: analogous.

*Markov: $X$ non-negative, $b > 0$: $\Pr[X \geq b] \leq \frac{\mathbb{E}[X]}{b}$.*

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**

\[
M_{X+Y}(t) = \mathbb{E}[e^{t(X+Y)}] = \mathbb{E}[e^{tX} \cdot e^{tY}] = \mathbb{E}[e^{tX}] \cdot \mathbb{E}[e^{tY}] = M_X(t) \cdot M_Y(t) \quad \checkmark
\]

**Concentration Inequality**

**Theorem (Chernoff Bounds):** Let \( X \) be a random variable and \( a > 0 \). Then, \( \Pr[X \geq a] \leq \min_{t > 0} \frac{\mathbb{E}[e^{tX}]/e^{ta}}{e^a} \) and \( \Pr[X \leq a] \leq \min_{t < 0} \frac{\mathbb{E}[e^{tX}]/e^{ta}}{e^a} \).

**Proof**

for all \( t > 0 \): \( \Pr[X \geq a] = \Pr[e^{tX} \geq e^{ta}] \leq \frac{\mathbb{E}[e^{tX}]/e^{ta}}{e^a} \leq \min_{t > 0} \frac{\mathbb{E}[e^{tX}]/e^{ta}}{e^a} \quad \checkmark 
\)

for all \( t < 0 \): analogous. \( \checkmark \)

Get bounds for specific random variables by finding a good \( t \)!

Looks scary, but is again just \( \mathbb{E}[f(X)] \) for \( f(X) = e^{tX} \)

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.

**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 $\epsilon > 0$

$$
\Pr[X \geq (1 + \epsilon)\mathbb{E}[X]] \leq \left(\frac{e^\epsilon}{(1 + \epsilon)^{1+\epsilon}}\right)^{\mathbb{E}[X]}.
$$

**Proof**

**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 $X_i \sim \text{Ber}(p)$.

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 $a > 0$:

$\Pr[X \geq a] \leq \min_{t>0} \mathbb{E}[e^{tX}] / e^{ta}$.

**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 $a > 0$:

$$\Pr[X \geq a] \leq \min_{t>0} \frac{\mathbb{E}[e^{tX}]}{e^{ta}}.$$

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 $X_i \sim \text{Ber}(p)$

$$M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t \cdot 0} + \Pr[X_i = 1] \cdot e^{t \cdot 1}$$

$$= (1 - p) + pe^t$$

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:** $M_X(t) = \mathbb{E}[e^{tX}]$
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 \( X_i \sim \text{Ber}(p) \)
\[
M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t \cdot 0} + \Pr[X_i = 1] \cdot e^{t \cdot 1}
= (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t - 1)p} \quad 1 + x \leq e^x
\]
\[
M_X(t) = M_{\sum X_i}(t) = \prod_{i=1}^{n} M_{X_i}(t)
\]

**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 $X_i \sim \text{Ber}(p)$

$$M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t \cdot 0} + \Pr[X_i = 1] \cdot e^{t \cdot 1} = (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t - 1)p}$$

$$M_X(t) = M_{\sum X_i}(t) = \prod_{i=1}^{n} M_{X_i}(t) \leq \prod_{i=1}^{n} e^{(e^t - 1)p}$$

**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 $\epsilon > 0$

$$
\Pr[X \geq (1 + \epsilon)\mathbb{E}[X]] \leq \left( \frac{e^\epsilon}{(1 + \epsilon)^{1+\epsilon}} \right)^n \mathbb{E}[X].
$$

**Proof** Consider $X$ as the sum of independent $X_i \sim \text{Ber}(p)$

$$
M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t \cdot 0} + \Pr[X_i = 1] \cdot e^{t \cdot 1}
= (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t - 1)p}
= e^{1 + x \leq e^x}
$$

$$
M_X(t) = M_\sum X_i(t) = \prod_{i=1}^n M_{X_i}(t) \leq \prod_{i=1}^n e^{(e^t - 1)p} = e^{(e^t - 1)np}
$$

| **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 $X_i \sim \text{Ber}(p)$

$$M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t\cdot 0} + \Pr[X_i = 1] \cdot e^{t\cdot 1} = (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t-1)p}.$$

$$M_X(t) = M_{\sum X_i}(t) = \prod_{i=1}^n M_{X_i}(t) \leq \prod_{i=1}^n e^{(e^t-1)p} = e^{(e^t-1)np} = e^{(e^t-1)\mathbb{E}[X]}.$$

- **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 $X_i \sim \text{Ber}(p)$

$$
M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t\cdot 0} + \Pr[X_i = 1] \cdot e^{t\cdot 1}
= (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t-1)p}
$$

$$
1 + x \leq e^x
$$

$$
M_X(t) = M\sum X_i(t) = \prod_{i=1}^n M_{X_i}(t) \leq \prod_{i=1}^n e^{(e^t-1)p} = e^{(e^t-1)\cdot np} = e^{(e^t-1)\cdot n\mathbb{E}[X]}
$$

$$
\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]]
$$

**Chernoff:** Random variable $X$ and $a > 0$:

$$
\Pr[X \geq a] \leq \min_{t>0} \frac{\mathbb{E}[e^{tX}]}{e^{ta}}.
$$

**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 $X \sim Bin(n, p)$. 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 $X_i \sim Ber(p)$

$$M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t \cdot 0} + \Pr[X_i = 1] \cdot e^{t \cdot 1} = (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t-1)p}$$

$$M_X(t) = M \sum X_i(t) = \prod_{i=1}^{n} M_{X_i}(t) \leq \prod_{i=1}^{n} e^{(e^t-1)p} = e^{(e^t-1)np} = e^{(e^t-1)\mathbb{E}[X]}$$

$$\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \min_{t>0} \frac{\mathbb{E}[e^{tx}]}{e^{t(1+\varepsilon)\mathbb{E}[X]}}$$

- **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:** $M_X(t) = \mathbb{E}[e^{tX}]$

- **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 $X_i \sim \text{Ber}(p)$

$$M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t \cdot 0} + \Pr[X_i = 1] \cdot e^{t \cdot 1} = (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t-1)p}$$

$$M_X(t) = M_{\sum X_i}(t) = \prod_{i=1}^n M_{X_i}(t) \leq \prod_{i=1}^n e^{(e^t-1)p} = e^{(e^t-1)np} = e^{(e^t-1)\mathbb{E}[X]}$$

$$\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \min_{t>0} \frac{\mathbb{E}[e^{tX}]}{e^{(1+\varepsilon)\mathbb{E}[X]}} \leq \min_{t>0} \frac{e^{(e^t-1)\mathbb{E}[X]}}{e^{(1+\varepsilon)\mathbb{E}[X]}}$$

**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 $X_i \sim \text{Ber}(p)$

$$M_X(t) = \mathbb{E}[e^{tX}] = \Pr[X_i = 0] \cdot e^{t\cdot 0} + \Pr[X_i = 1] \cdot e^{t\cdot 1} = (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t - 1)p}$$

$$M_X(t) = M_{\sum X_i}(t) = \prod_{i=1}^{n} M_{X_i}(t) \leq \prod_{i=1}^{n} e^{(e^t - 1)p} = e^{(e^t - 1)np} = e^{(e^t - 1)\mathbb{E}[X]}$$

$$\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \min_{t > 0} \frac{\mathbb{E}[e^{tX}]^{e\varepsilon}}{et(1+\varepsilon)\mathbb{E}[X]} \leq \min_{t > 0} \frac{e^{(e^t - 1)\mathbb{E}[X]}}{et(1+\varepsilon)\mathbb{E}[X]} = \min_{t > 0} \left( \frac{e^{(e^t - 1)}}{et(1+\varepsilon)} \right)^{\mathbb{E}[X]}$$
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 \( X_i \sim \text{Ber}(p) \)
\[
M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t\cdot0} + \Pr[X_i = 1] \cdot e^{t\cdot1}
= (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t - 1)p}
\]
\[
1 + x \leq e^x
\]
\[
M_X(t) = \prod_{i=1}^{n} M_{X_i}(t) \leq \prod_{i=1}^{n} e^{(e^t - 1)p} = e^{(e^t - 1)np}
= e^{(e^t - 1)\mathbb{E}[X]}
\]
\[
\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \min_{t > 0} \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\varepsilon)\mathbb{E}[X]}} \leq \min_{t > 0} \frac{e^{(e^t - 1)\mathbb{E}[X]}}{e^{t(1+\varepsilon)\mathbb{E}[X]}} = \min_{t > 0} \left( \frac{e^\varepsilon}{e^{t(1+\varepsilon)}} \right)^{\mathbb{E}[X]} \leq \left( \frac{e^\varepsilon}{(1+\varepsilon)^{1+\varepsilon}} \right)^{\mathbb{E}[X]}
\]
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 $X_i \sim \text{Ber}(p)$

$$M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t \cdot 0} + \Pr[X_i = 1] \cdot e^{t \cdot 1} = (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t - 1)p} \leq e^{(1 + x)e^t}$$

$$M_X(t) = M_{\sum X_i}(t) = \prod_{i=1}^{n} M_{X_i}(t) \leq \prod_{i=1}^{n} e^{(e^t - 1)p} = e^{(e^t - 1)np} = e^{(e^t - 1)\mathbb{E}[X]}$$

$$\Pr[X \geq (1 + \varepsilon) \mathbb{E}[X]] \leq \min_{t > 0} \frac{\mathbb{E}[e^{tX}]^{\mathbb{E}[X]}}{e^{t(1+\varepsilon)\mathbb{E}[X]}} \leq \min_{t > 0} \frac{e^{(e^t - 1)\mathbb{E}[X]}}{e^{t(1+\varepsilon)\mathbb{E}[X]}} = \min_{t > 0} \left( \frac{e^{(e^t - 1)}}{e^{t(1+\varepsilon)}} \right)^{\mathbb{E}[X]} \leq \left( \frac{e^\varepsilon}{(1 + \varepsilon)^{1 + \varepsilon}} \right)^{\mathbb{E}[X]}$$

**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 $X_i \sim \text{Ber}(p)$

$$M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t \cdot 0} + \Pr[X_i = 1] \cdot e^{t \cdot 1} = (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t - 1)p} \quad 1 + x \leq e^x$$

$$M_X(t) = M_X(t) = \prod_{i=1}^{n} M_{X_i}(t) \leq \prod_{i=1}^{n} e^{(e^t - 1)p} = e^{(e^t - 1)np} = e^{(e^t - 1)\mathbb{E}[X]}$$

$$\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \min_{t > 0} \frac{\mathbb{E}[e^{tX}]}{e^{t(1 + \varepsilon)\mathbb{E}[X]}} \leq \min_{t > 0} \frac{e^{(e^t - 1)\mathbb{E}[X]}}{e^{t(1 + \varepsilon)\mathbb{E}[X]}} = \min_{t > 0} \left(\frac{e^{e^t - 1}}{e^{t(1 + \varepsilon)}}\right)^{\mathbb{E}[X]} \leq \left(\frac{e^\varepsilon}{(1 + \varepsilon)^{(1 + \varepsilon)}}\right)^{\mathbb{E}[X]}$$

**Example**

- Sum of 20 unfair $\{0, 1\}$-coin tosses: $X \sim \text{Bin}(20, \frac{1}{5})$, $\mathbb{E}[X] = 4$
- $\Pr[X \geq 16] = \Pr[X \geq (1 + 3)\mathbb{E}[X]]$

**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 $X_i \sim \text{Ber}(p)$

$$
M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t\cdot 0} + \Pr[X_i = 1] \cdot e^{t\cdot 1}
= (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t - 1)p}
$$

$$
M_X(t) = M\sum X_i(t) = \prod_{i=1}^{n} M_{X_i}(t) \leq \prod_{i=1}^{n} e^{(e^t - 1)p} = e^{(e^t - 1)p n}
= e^{(e^t - 1) \mathbb{E}[X]}
$$

$$
\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \min_{t > 0} \frac{\mathbb{E}[e^{tX}]}{e^{t(1 + \varepsilon)\mathbb{E}[X]}} \leq \min_{t > 0} \frac{e^{(e^t - 1)\mathbb{E}[X]}}{e^{t(1 + \varepsilon)\mathbb{E}[X]}} = \min_{t > 0} \left(\frac{e^{(e^t - 1)}}{e^{t(1 + \varepsilon)}}\right)^\mathbb{E}[X] \leq \left(\frac{e^\varepsilon}{(1 + \varepsilon)^{(1 + \varepsilon)}}\right)^\mathbb{E}[X]
$$

**Example**
- Sum of 20 unfair $\{0, 1\}$-coin tosses: $X \sim \text{Bin}(20, \frac{1}{5})$, $\mathbb{E}[X] = 4$
- $\Pr[X \geq 16] = \Pr[X \geq (1 + 3)\mathbb{E}[X]] \leq \left(\frac{e^{3}}{(1 + 3)^{(1 + 3)}}\right)^4$
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 $X_i \sim \text{Ber}(p)$

$$M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t \cdot 0} + \Pr[X_i = 1] \cdot e^{t \cdot 1} = (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t - 1)p} = e^{(1 + x) \cdot np} = e^{(e^t - 1)\mathbb{E}[X]}$$

$$M_X(t) = M\sum_{i=1}^n X_i(t) = \prod_{i=1}^n M_{X_i}(t) \leq \prod_{i=1}^n e^{(e^{t-1})p} = e^{(e^{t-1})n\mathbb{E}[X]} = e^{(e^{t-1})\mathbb{E}[X]}$$

$$\Pr[X \geq (1 + \epsilon)\mathbb{E}[X]] \leq \min_{t > 0} \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\epsilon)\mathbb{E}[X]]}} \leq \min_{t > 0} \frac{e^{(e^{t-1})\mathbb{E}[X]}}{e^{t(1+\epsilon)\mathbb{E}[X]}} = \min_{t > 0} \left(\frac{e^{e^{t-1}}}{e^{t(1+\epsilon)}}\right)^{\mathbb{E}[X]} \leq \left(\frac{e^\epsilon}{(1 + \epsilon)^{(1+\epsilon)}}\right)^{\mathbb{E}[X]}$$

**Example**

- Sum of 20 unfair {0, 1}-coin tosses: $X \sim \text{Bin}(20, \frac{1}{5})$, $\mathbb{E}[X] = 4$

- $\Pr[X \geq 16] = \Pr[X \geq (1 + 3)\mathbb{E}[X]] \leq \left(\frac{e^3}{(1+3)^{1+3}}\right)^4 = \frac{e^{12}}{4^4} \approx 0.00003789$

**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 $X, Y$:

$$M_{X+Y}(t) = M_X(t) \cdot M_Y(t).$$
Application: Binomial Distribution

**Theorem:** Let $X \sim \text{Bin}(n, p)$. Then for any $\varepsilon > 0$

$$\Pr[X \geq (1 + \varepsilon)E[X]] \leq \left( \frac{e^\varepsilon}{(1 + \varepsilon)^{(1+\varepsilon)}} \right)^{E[X]}.$$

**Proof** Consider $X$ as the sum of independent $X_i \sim \text{Ber}(p)$

$$M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t \cdot 0} + \Pr[X_i = 1] \cdot e^{t \cdot 1}$$

$$= (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t - 1)p}$$

$$1 + x \leq e^x$$

$$M_X(t) = M_{\sum X_i}(t) = \prod_{i=1}^{n} M_{X_i}(t) \leq \prod_{i=1}^{n} e^{(e^t - 1)p} = e^{(e^t - 1)np} = e^{(e^t - 1)E[X]}$$

$$\Pr[X \geq (1 + \varepsilon)E[X]] \leq \min_{t > 0} t > 0 \mathbb{E}[e^{tX}] \leq \min_{t > 0} e^{(e^t - 1)E[X]} = \min_{t > 0} \left( e^{(e^t - 1)E[X]} \right)^{E[X]} \leq \left( \frac{e^\varepsilon}{(1 + \varepsilon)^{(1+\varepsilon)}} \right)^{E[X]}$$

**Example**

- Sum of 20 unfair $\{0, 1\}$-coin tosses: $X \sim \text{Bin}(20, \frac{1}{5})$, $E[X] = 4$
- $\Pr[X \geq 16] = \Pr[X \geq (1 + 3)E[X]] \leq \left( \frac{e^{3}}{(1+3)^{1+3}} \right)^{4} = \frac{e^{12}}{4^{4}} \approx 0.00003789$

**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:** $\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$

**Chebychev:** $\Pr[X \geq (1 + \varepsilon)E[X]] \leq \frac{\mathbb{E}[X]}{(1 + \varepsilon)^{(1+\varepsilon)}}$

**Actual:** $\approx 0.0000000138$
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 $\varepsilon > 0$

$$\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \left(\frac{e^\varepsilon}{(1 + \varepsilon)^{1+\varepsilon}}\right)^{\mathbb{E}[X]}.$$

**Chernoff:** Random variable $X$ and $a > 0$:

$$\Pr[X \geq a] \leq \min_{t>0} \frac{\mathbb{E}[e^{tX}]}{e^{ta}}.$$

**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 $\varepsilon > 0$
$$\Pr[X \geq (1 + \varepsilon)E[X]] \leq \left(\frac{e^\varepsilon}{(1 + \varepsilon)^{1+\varepsilon}}\right)^{E[X]}.$$  

**Corollary:** Let $X \sim \text{Bin}(n, p)$. Then for any $t \geq 6E[X]$, $\Pr[X \geq t] \leq 2^{-t}$.

**Corollary:** Let $X \sim \text{Bin}(n, p)$. Then for any $\varepsilon \in (0, 1]$, $\Pr[X \geq (1 + \varepsilon)E[X]] \leq e^{-\varepsilon^2 / 3 \cdot E[X]}$.  

**Chernoff:** Random variable $X$ and $a > 0$:
$$\Pr[X \geq a] \leq \min_{t > 0} \frac{E[e^{tX}]}{e^{ta}}.$$
Chernoff – Simpler Versions

**Theorem**: Let $X \sim \text{Bin}(n, p)$. Then for any $\varepsilon > 0$
\[ \Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \left( \frac{e^\varepsilon}{(1 + \varepsilon)^{1+\varepsilon}} \right)^{\mathbb{E}[X]} . \]

**Chernoff**: Random variable $X$ and $a > 0$:
\[ \Pr[X \geq a] \leq \min_{t>0} \mathbb{E}[e^{tX}] / e^{ta} . \]

**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 $\varepsilon \in (0, 1)$, $\Pr[X \geq (1 - \varepsilon)\mathbb{E}[X]] \leq e^{-\varepsilon^2/2\cdot\mathbb{E}[X]}$. 
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 $\varepsilon \in (0, 1]$, $\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq e^{-\varepsilon^2/3}\mathbb{E}[X]$.

**Corollary**: Let $X \sim \text{Bin}(n, p)$. Then for any $\varepsilon \in (0, 1)$, $\Pr[X \geq (1 - \varepsilon)\mathbb{E}[X]] \leq e^{-\varepsilon^2/2}\mathbb{E}[X]$.

In fact, these also work when the $X_i$ are Bernoulli random variables with different success probabilities.
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