CSE 312

Foundations of Computing II

Lecture 21: Chernoff Bound & Union Bound

Review Tail Bounds

Putting a limit on the probability that a random variable is in the "tails" of the distribution (e.g., not near the middle).

Usually statements in the form of

$$P(X \geq a) \leq b$$

or

$$P(|X - \mathbb{E}[X]| \geq a) \leq b$$

Review Markov's and Chebyshev's Inequalities

Theorem (Markov's Inequality). Let X be a random variable taking only non-negative values. Then, for any $t > 0$,

$$P(X \geq t) \leq \frac{\mathbb{E}[X]}{t}.$$

$P(X \geq k \cdot E[X]) \leq \frac{1}{k}$

Theorem (Chebyshev's Inequality). Let X be a random variable. Then, for any $t > 0$,

$$P(|X - \mathbb{E}[X]| \geq t) \leq \frac{\mathrm{Var}(X)}{t^2}.$$

Agenda

- Chernoff Bound
- Example: Server Load, and the union bound

Chebyshev & Binomial

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq t) \leq \frac{\mathrm{Var}(X)}{t^2}.$$

Reformulated: $P(|X - \mu| \geq \delta\mu) \leq \frac{\sigma^2}{\delta^2\mu^2}$ where $\mu = \mathbb{E}[X]$ and $\sigma^2 = \mathrm{Var}(X)$

t t^2

If $X \sim \mathrm{Bin}(n, p)$, then $\mu = np$ and $\sigma^2 = np(1-p)$

$$P(|X - \underbrace{\mu}_{np}| \geq \delta\mu) \leq \frac{np(1-p)}{\delta^2 n^2 p^2} = \frac{1}{\delta^2} \cdot \frac{1}{n} \cdot \frac{1-p}{p}$$

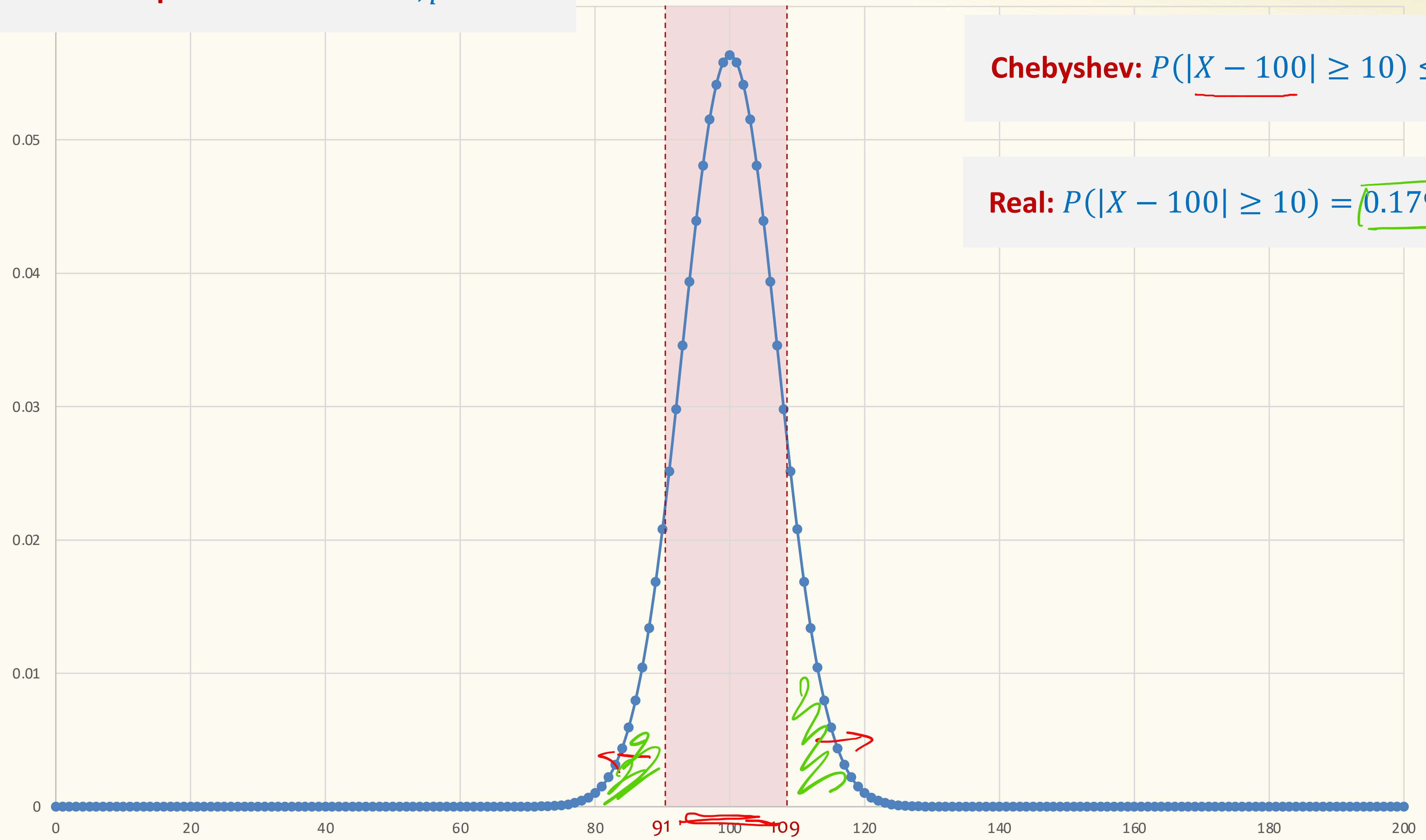
$\frac{1-p}{p} = 1$ $\quad \frac{1}{\delta^2} = 100$

E.g., $\delta = 0.1$, $p = 0.5$: $n = 200$: $P(|X - \mu| \geq \delta\mu) \leq 0.5$ $\;= \frac{100}{n}$

$n = 800$: $P(|X - \mu| \geq \delta\mu) \leq 0.125$

How good is it?

Binomial with parameter $n = 200, p = 0.5$



Chebyshev: $P(|X - 100| \geq 10) \leq \frac{1}{2}$

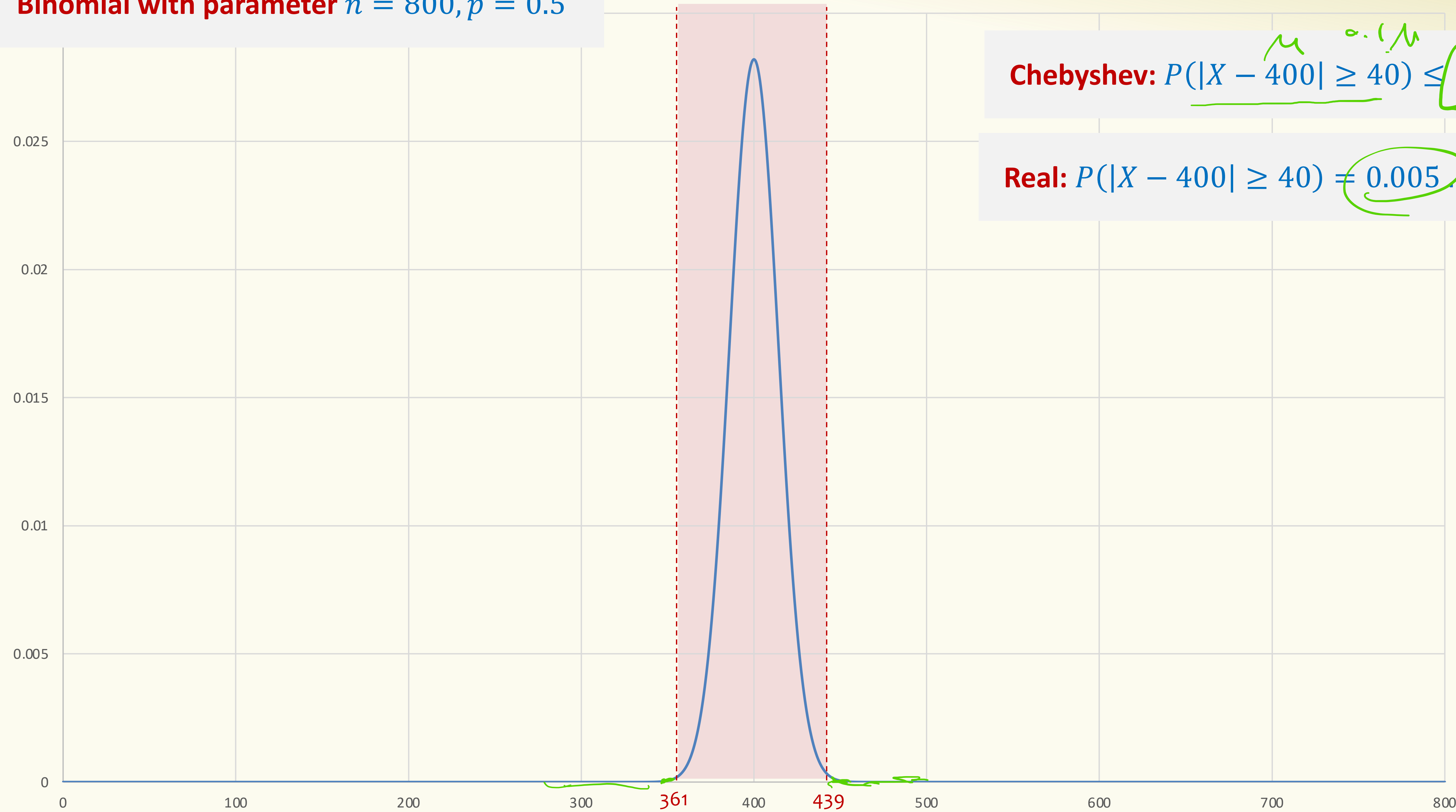
Real: $P(|X - 100| \geq 10) = \boxed{0.179 \ldots}$

PMF

Binomial with parameter $n = 800, p = 0.5$

Chebyshev: $P(|X - 400| \geq 40) \leq 0.125$

Real: $P(|X - 400| \geq 40) = 0.005 \ldots$



Chernoff-Hoeffding Bound

Theorem. Let $X = X_1 + \cdots + X_n$ be a sum of independent RVs, each taking values in $[0,1]$, such that $\mathbb{E}[X] = \mu$. Then, for every $\delta \in [0,1]$,

$$P(|X - \mu| \geq \delta \cdot \mu) \leq e^{-\frac{\delta^2 \mu}{4}}.$$

$\Theta(\delta^2 \mu)$

$2 \leq 2 / \ldots$

Herman Chernoff, Herman Rubin, Wassily Hoeffding

Example: If $X \sim \mathrm{Bin}(n, p)$, then $X = X_1 + \cdots + X_n$ is a sum of independent $\{0,1\}$-Bernoulli variables, and $\mu = np$

Note: More accurate versions are possible, but with more cumbersome right-hand side (e.g., see textbook)

Chernoff-Hoeffding Bound – Binomial Distribution

Theorem. (CH bound, binomial case) Let $X \sim \mathrm{Bin}(n, p)$. Let $\mu = np = \mathbb{E}[X]$. Then, for any $\delta \in [0,1]$,

$$P(|X - \mu| \geq \delta \cdot \mu) \leq e^{-\frac{\delta^2 np}{4}}.$$

Example:

$p = 0.5$
$\delta = 0.1$

	Chebyshev	Chernoff
n	$\frac{1}{\delta^2} \cdot \frac{1}{n} \cdot \frac{1-p}{p}$	$e^{-\frac{\delta^2 np}{4}}$
800	0.125	0.3679
2600	0.03846	0.03877
8000	0.0125	0.00005
80000	0.00125	3.72 $\times$ 10^-44^

Chernoff Bound – Example

$$\mathbb{P}(|X - \mu| \geq \delta \cdot \mu) \leq e^{-\frac{\delta^2 np}{4}}.$$

Alice tosses a fair coin n times, what is an upper bound for the probability that she sees heads at least $0.75 \times n$ times?

X = # heads $P(X \geq 0.75n)$ $X \sim Bin(n, 1/2)$

$\mu = n/2$

$= P(X - \mu \geq 0.25n)$

$\leq P(|X - \mu| \geq 0.25n)$ → 0.5

$\delta\mu$

$\leq e^{-\frac{\delta^2 n/2}{4}}$

Poll: pollev.com/stefanotessaro617

a. $e^{-n/64}$ 2

b. $e^{-n/32}$ 10

c. $e^{-n/16}$ 1

d. $e^{-n/8}$ 4

Chernoff vs Chebyshev – Summary

$$\frac{1}{\delta^2} \cdot \frac{1}{n} \cdot \frac{1-p}{p}$$

Chebyshev, <u>linear</u> decrease in n

vs

Chernoff, <u>exponential</u> decrease in n

$$e^{-\frac{\delta^2 np}{4}}$$

Why is the Chernoff Bound True?

Proof strategy (upper tail): For any $t > 0$:

- $P(X \geq (1+\delta)\cdot\mu) = P\left(e^{tX} \geq e^{t(1+\delta)\cdot\mu}\right)$
- Then, apply Markov + independence:

$$P\left(e^{tX} \geq e^{t(1+\delta)\cdot\mu}\right) \leq \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\delta)\mu}} = \frac{\mathbb{E}[e^{tX_1}]\cdots\mathbb{E}[e^{tX_n}]}{e^{t(1+\delta)\mu}}$$

- Find t minimizing the right-hand-side.

Brain Break

Agenda

- Chernoff Bound
- Example: Server Load, and the union bound

Application – Distributed Load Balancing

We have k processors, and $n \gg k$ jobs.
We want to distribute jobs evenly across processors.

Strategy: Each job assigned to a ~~randomly~~ uniformly chosen processor!

X_i = load of processor i $\quad X_i \sim \mathrm{Binomial}(n, 1/k) \quad \mathbb{E}[X_i] = n/k$

$X = \max\{X_1, \dots, X_k\}$ = max load of a processor

Question: How close is X to n/k?

Distributed Load Balancing

Claim. (Load of single server) If $n > 16k \ln k$, then

$$P\left(X_i > \frac{n}{k} + 4\boxed{\sqrt{\frac{n \ln k}{k}}}\right) \leq 1/k^4.$$

Example:

- $n = 10^6 \gg k = 1000$
- $\frac{n}{k} + 4\sqrt{n \ln k / k} \approx 1332$
- *"The probability that server i processes more than 1332 jobs is at most 1-over-one-trillion!"*

Distributed Load Balancing

Claim. (Load of single server) If $n > 16k \ln k$, then

$$P\left(X_i > \underbrace{\frac{n}{k}}_{\mu} + 4\sqrt{\frac{n \ln k}{k}}\right) = P\left(X_i > \frac{n}{k}\left(\underbrace{1 + 4\sqrt{\frac{k \ln k}{n}}}_{1+\delta}\right)\right) \le 1/k^4.$$

δ

Proof. Set $\mu = \mathbb{E}[X_i] = \frac{n}{k}$ and $\delta = 4\sqrt{\frac{k}{n}\ln k} < 4\sqrt{\frac{k}{16k \ln k}\ln k} = 1$

$n > 16k \ln k$

$$P\left(X_i > \mu\left(1 + 4\sqrt{\frac{k \ln k}{n}}\right)\right) = P\big(X_i > \mu(1+\delta)\big)$$

$$\le P\big(|X_i - \mu| > \mu\delta\big)$$

$$\le e^{-\frac{\delta^2 \mu}{4}} = e^{-4 \ln k} = \frac{1}{k^4}$$

What about the maximum load?

Claim. (Load of single server) If $n > 16k \ln k$, then

$$P\left(X_i > \frac{n}{k} + 4\sqrt{\frac{n \ln k}{k}}\right) \leq 1/k^4.$$

What about $X = \max\{X_1, \ldots, X_k\}$?

Note: $X_1, \ldots, X_k$ are <u>not</u> (mutually) independent!

In particular: $X_1 + \cdots + X_k = n$

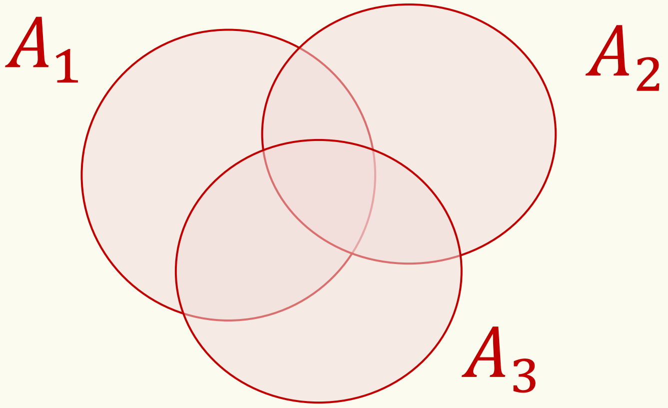
When non-trivial outcome of one RV can be derived from other RVs, they are non-independent.

Detour – Union Bound

Theorem (Union Bound). Let $A_1, \ldots, A_n$ be arbitrary events. Then,

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i)$$

Intuition (3 evts.):



Detour – Union Bound - Example

Suppose we have $N = 200$ computers, where each one fails with probability 0.001.

What is the probability that at least one server fails?

Let A_i be the event that server i fails.

Then event that at least one server fails is $\cup_{i=1}^{N} A_i$

$$P\left(\bigcup_{i=1}^{N} A_i\right) \le \sum_{i=1}^{N} P(A_i) = 0.001N = 0.2$$

What about the maximum load?

Claim. (Load of single server) If $n > 16k \ln k$, then

$$P\left(X_i > \frac{n}{k} + 4\sqrt{\frac{n \ln k}{k}}\right) \leq 1/k^4.$$

What about $X = \max\{X_1, \dots, X_k\}$?

$$P\left(X > \frac{n}{k} + 4\sqrt{n \ln k / k}\right) = P\left(\left\{X_1 > \frac{n}{k} + 4\sqrt{n \ln k / k}\right\} \cup \cdots \cup \left\{X_k > \frac{n}{k} + 4\sqrt{n \ln k / k}\right\}\right)$$

Union bound

$$\leq P\left(X_1 > \frac{n}{k} + 4\sqrt{\frac{n \ln k}{k}}\right) + \cdots + P\left(X_k > \frac{n}{k} + 4\sqrt{n \ln k / k}\right)$$

$$\leq \frac{1}{k^4} + \cdots + \frac{1}{k^4} = k \times \frac{1}{k^4} = \frac{1}{k^3}$$

What about the maximum load?

Claim. (Load of single server) If $n > 16k \ln k$, then

$$P\left(X_i > \frac{n}{k} + 4\sqrt{\frac{n \ln k}{k}}\right) \leq 1/k^4.$$

Claim. (Max load) Let $X = \max\{X_1, \dots, X_k\}$. If $n > 16k \ln k$, then

$$P\left(X > \frac{n}{k} + 4\sqrt{\frac{n \ln k}{k}}\right) \leq 1/k^3. = k \cdot \frac{1}{k^4}$$