CSE 312

Foundations of Computing II

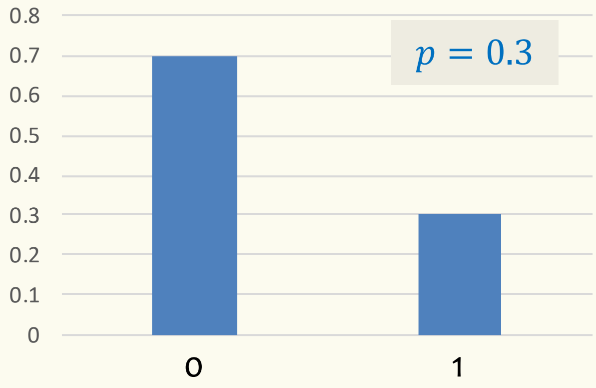
Lecture 17: Poisson Distribution + Chernoff Intro

PAUL G. ALLEN SCHOOL
OF COMPUTER SCIENCE & ENGINEERING

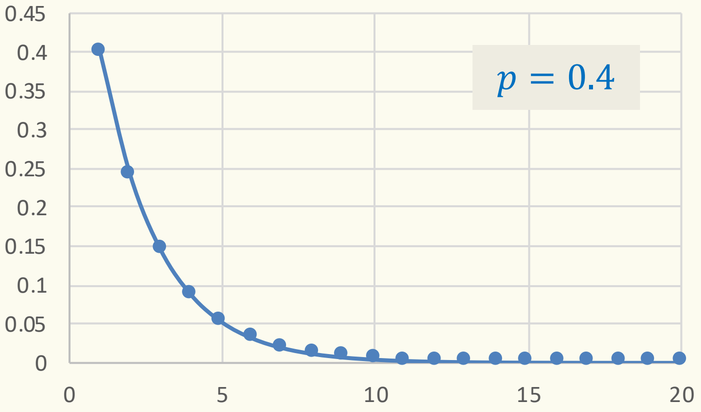
Stefano Tessaro
tessaro@cs.washington.edu

Distributions – Recap

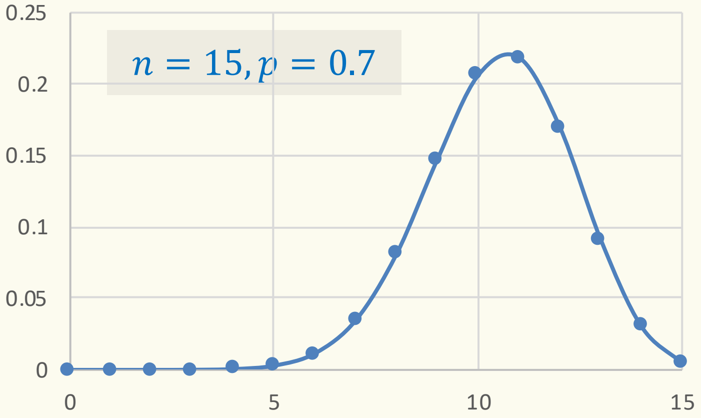
Name	Pars	Range	PMF	Expectation	Variance
Bernoulli	p	$\{0,1\}$	$\mathbb{p}(1) = p,\ \mathbb{p}(0) = 1 - p$	p	$p(1-p)$
Geometric	p	$\{1,2,3,\dots\} = \mathbb{N}^+$	$\mathbb{p}(i) = (1-p)^{i-1}p$	$1/p$	$(1-p)/p^2$
Binomial	n, p	$\{0,1,\dots,n\}$	$\mathbb{p}(k) = \binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$



Bernoulli



Geometric

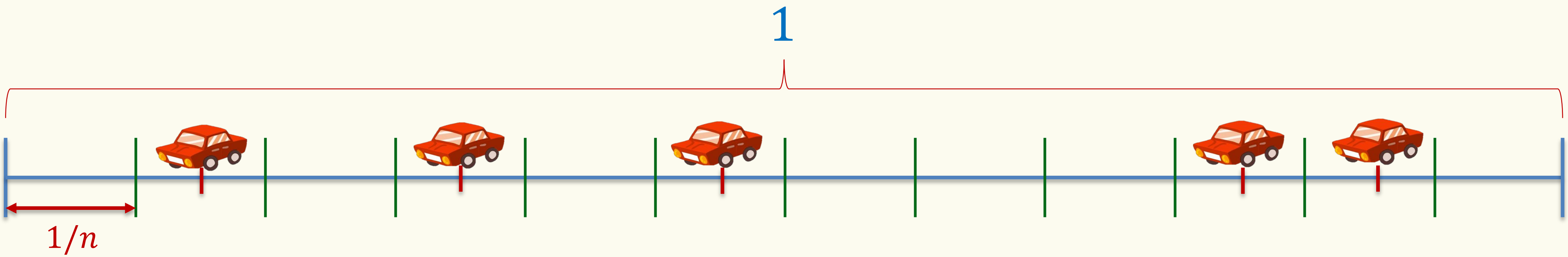


Binomial

Example – Number of Cars

X = # cars passing through an intersection in 1 hour

Wanted: $\mathbb{E}(X) = \lambda$ for some given $\lambda > 0$



Discretize problem: n intervals, each of length $\frac{1}{n}$.

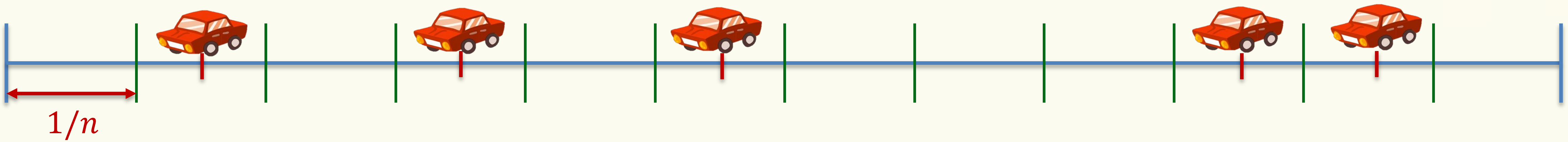
Bernoulli $X_i = 1$ if car in i-th interval (0 otherwise). $\mathbb{P}(X_i = 1) = \frac{\lambda}{n}$

$X = \sum_{i=1}^{n} X_i$

X is <u>binomial</u> $\mathbb{P}(X = i) = \binom{n}{i}\left(\frac{\lambda}{n}\right)^i\left(1 - \frac{\lambda}{n}\right)^{n-i}$

Motivation – Number of Cars

X is <u>binomial</u> $\mathbb{P}(X = i) = \binom{n}{i}\left(\frac{\lambda}{n}\right)^i\left(1 - \frac{\lambda}{n}\right)^{n-i}$



We want now $n \to \infty$

$$\mathbb{P}(X = i) = \binom{n}{i}\left(\frac{\lambda}{n}\right)^i\left(1 - \frac{\lambda}{n}\right)^{n-i} = \underbrace{\frac{n!}{(n-i)!}}_{\to 1}\frac{\lambda^i}{i!}\underbrace{\left(1 - \frac{\lambda}{n}\right)^n}_{\to e^{-\lambda}}\underbrace{\left(1 - \frac{\lambda}{n}\right)^{-i}}_{\to 1}$$

Poisson Random Variables

Siméon Denis Poisson
1781-1840

Definition. A **Poisson random variable** X with parameter $\lambda \geq 0$ is such that for all $i = 0,1,2,3 \ldots,$

$$\mathbb{P}(X = i) = e^{-\lambda} \cdot \frac{\lambda^i}{i!}$$

Several examples of "Poisson processes":

- # of requests to web servers in an hour
- # of photons hitting a light detector in a given interval
- # of patients arriving to ER within an hour
- …

General principle: Infinitely small interval, counting # of occurrences of event, each individual event can happen (at most once) with same chance in every interval.

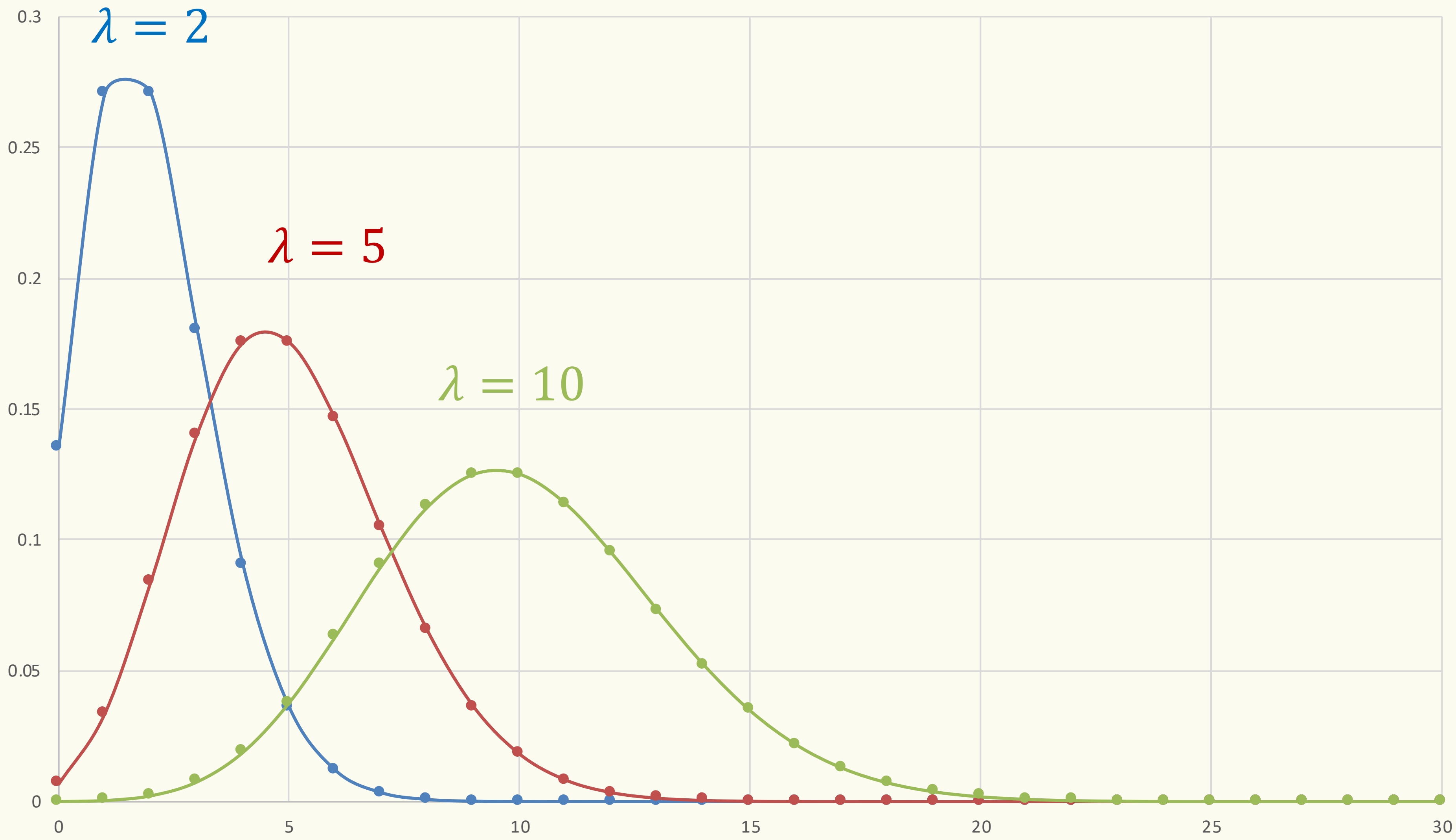
Validity of Distribution

We first want to verify that Possion probabilities sum up to 1.

$$\sum_{i=0}^{\infty} \mathbb{P}(X = i) = e^{-\lambda} \underbrace{\sum_{i=0}^{\infty} \frac{\lambda^i}{i!}} = e^{-\lambda} e^{\lambda} = 1$$

Fact. $\sum_{i=0}^{\infty} \frac{x^i}{i!} = e^x$

Probability Mass Function

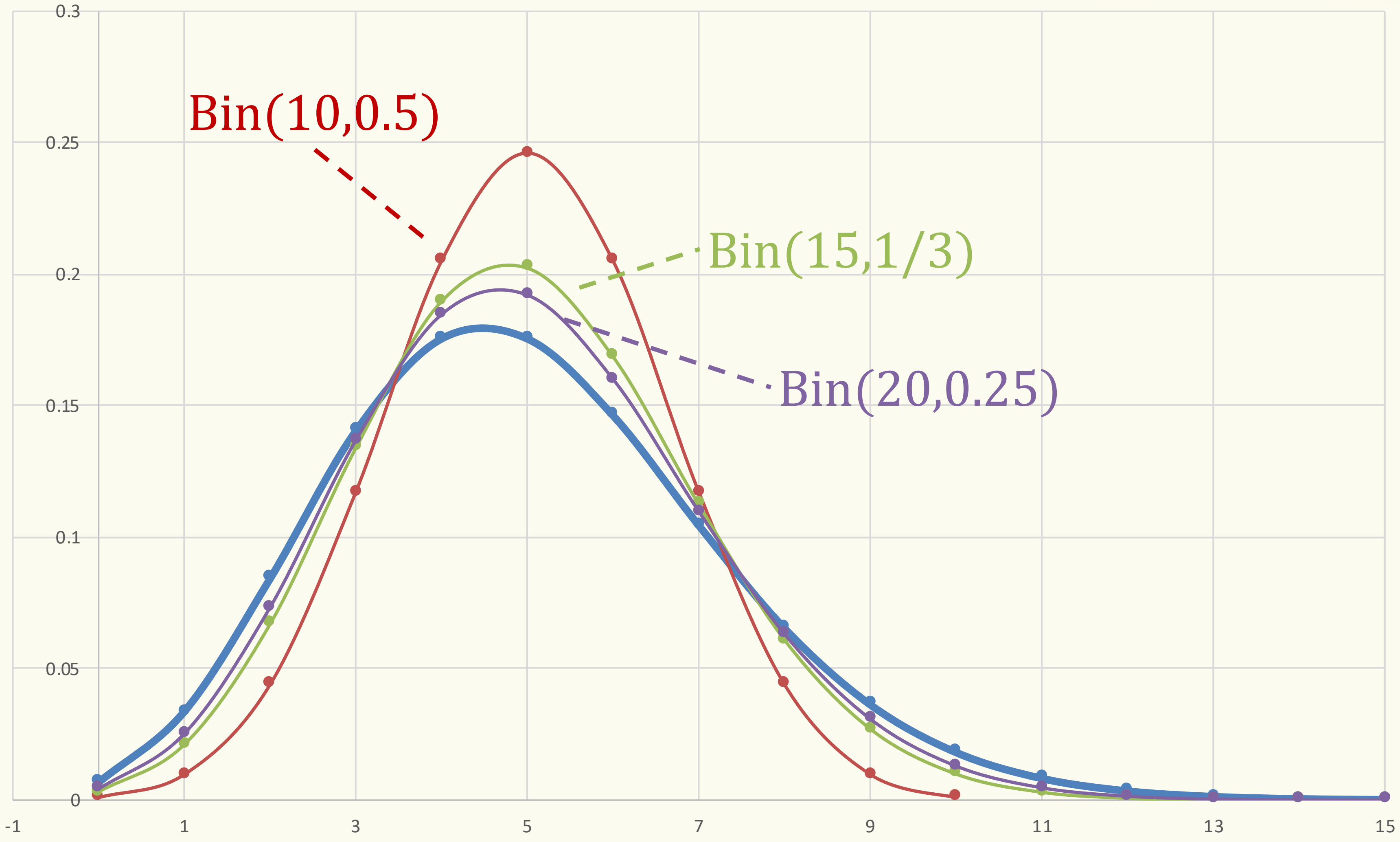


Probability Mass Function – Convergence of Binomials

$\lambda = 5$

$p = \frac{5}{n}$

$n = 10{,}15{,}20$



Expectation

We know this <u>by design</u> (limit of Binomial with expectation λ), but formally, this needs a proof.

Theorem. If X is a Poisson RV with parameter λ, then

$$\mathbb{E}(X) = \lambda$$

Proof.

$$\mathbb{E}(X) = \sum_{i=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^i}{i!} \cdot i = \sum_{i=1}^{\infty} e^{-\lambda} \cdot \frac{\lambda^i}{(i-1)!}$$

$$= \lambda \sum_{i=1}^{\infty} e^{-\lambda} \cdot \frac{\lambda^{i-1}}{(i-1)!}$$

= 1 (see prior slides!)

$$= \lambda \sum_{i=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^i}{i!} = \lambda \cdot 1 = \lambda$$

Variance

Intuitively (limit of Binomial): $\mathrm{Var}(X) = np(1-p) = \frac{n\lambda}{n}\left(1-\frac{\lambda}{n}\right) = \lambda\left(1-\frac{\lambda}{n}\right) \to \lambda$

Theorem. If X is a Poisson RV with parameter λ, then $\mathrm{Var}(X) = \lambda$

Proof. $\mathbb{E}(X^2) = \sum_{i=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^i}{i!} \cdot i^2 = \sum_{i=1}^{\infty} e^{-\lambda} \cdot \frac{\lambda^i}{(i-1)!} i$

$$= \lambda \sum_{i=1}^{\infty} e^{-\lambda} \cdot \frac{\lambda^{i-1}}{(i-1)!} \cdot i \quad = \lambda \sum_{i=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^i}{i!} \cdot (i+1)$$

$$= \lambda \left[\underbrace{\sum_{i=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^i}{i!} \cdot i}_{= \mathbb{E}(X) = \lambda} + \underbrace{\sum_{i=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^i}{i!}}_{= 1} \right] = \lambda^2 + \lambda$$

Variance – Proof (cont'd)

Theorem. If X is a Poisson RV with parameter λ, then $\mathrm{Var}(X) = \lambda$

We now know: $\mathbb{E}(X^2) = \lambda^2 + \lambda$

$\mathbb{E}(X) = \lambda$

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

Discrete Distributions – Final Recap

Name	Pars	Range	PMF	Exp.	Var.
Bernoulli	p	$\{0,1\}$	$\mathbb{p}(1) = p$, $\mathbb{p}(0) = 1 - p$	p	$p(1-p)$
Geometric	p	$\{1,2,3,\ldots\} = \mathbb{N}^+$	$\mathbb{p}(i) = (1-p)^{i-1}p$	$1/p$	$(1-p)/p^2$
Binomial	n, p	$\{0,1,\ldots,n\}$	$\mathbb{p}(k) = \binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$
Poisson	λ	$\{0,1,2,3,\ldots\} = \mathbb{N}$	$\mathbb{p}(i) = e^{-\lambda} \cdot \frac{\lambda^i}{i!}$	λ	λ

Other common distributions: Hypergeometric (see Quiz Section), negative binomials (= sum of geometric)

0 1

Next – Concentration and its applications

General question: How close is a random variable to its expectation?

So far: Markov's inequality + Chebyshev's inequality
[Also cf. HW5 + Section 6]

Example

Flip n independent coins, each heads with probability p

X = # of flips which are heads

We know that X is binomial: $\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$

$$\mathbb{E}(X) = n \cdot p$$

$$\mathrm{Var}(X) = n \cdot p \cdot (1 - p)$$

Question: What is the probability that X is within 10% of the expectation?

Deviation via Chebyshev

Theorem. Let X be a random variable. Then, for any $t > 0$,

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

Use Chebyshev's inequality

$$\mathbb{P}\big(|X - \mathbb{E}(X)| \geq \epsilon \cdot \mathbb{E}(X)\big) \leq \frac{np(1-p)}{\epsilon^2 n^2 p^2} = \frac{1-p}{\epsilon^2 np}$$

E.g. $\epsilon = 0.1, p = 0.5$

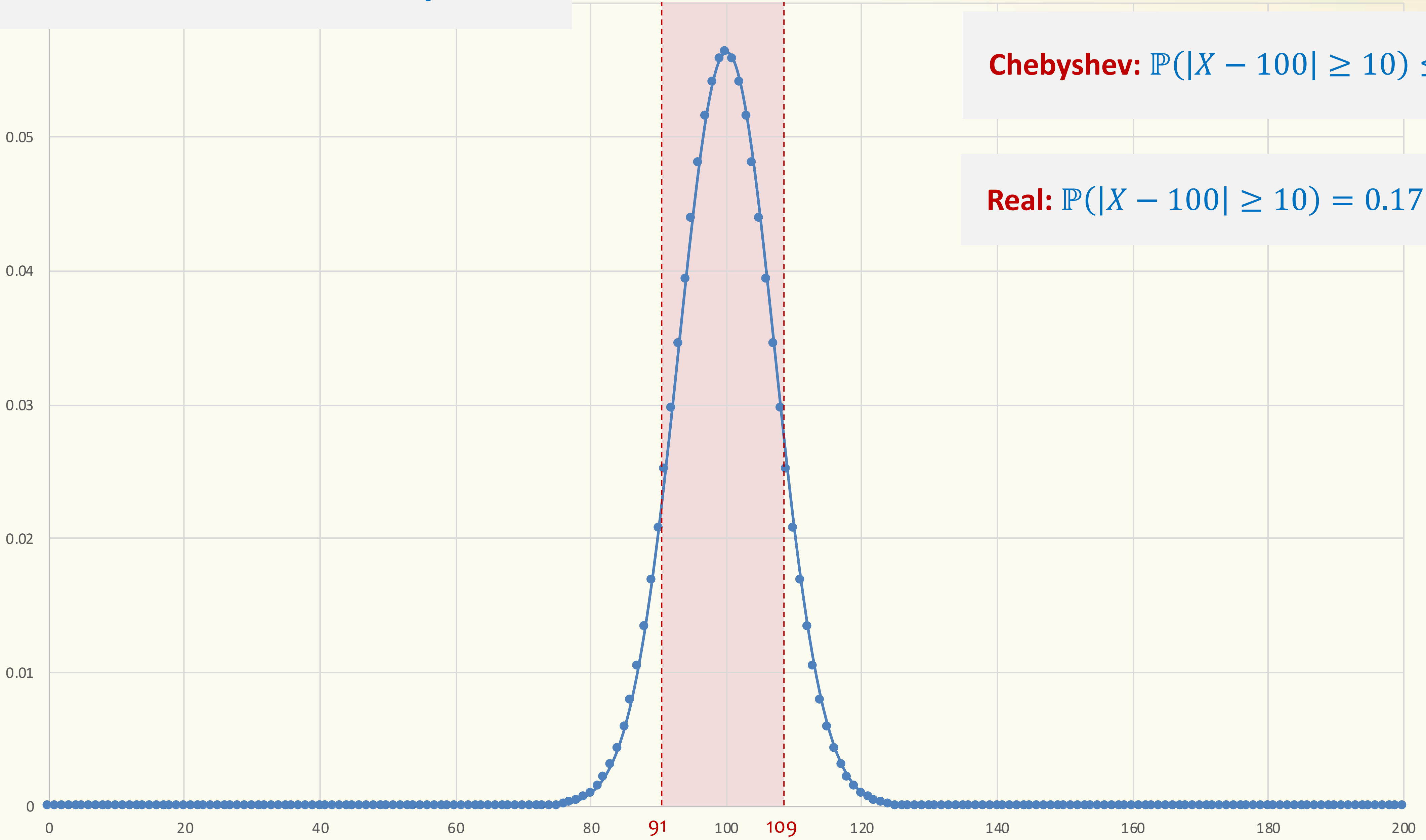
$$\mathbb{P}\big(|X - \mathbb{E}(X)| \geq 0.1 \cdot \mathbb{E}(X)\big) \leq \frac{100}{n} \to 0$$

Is this a good estimate?

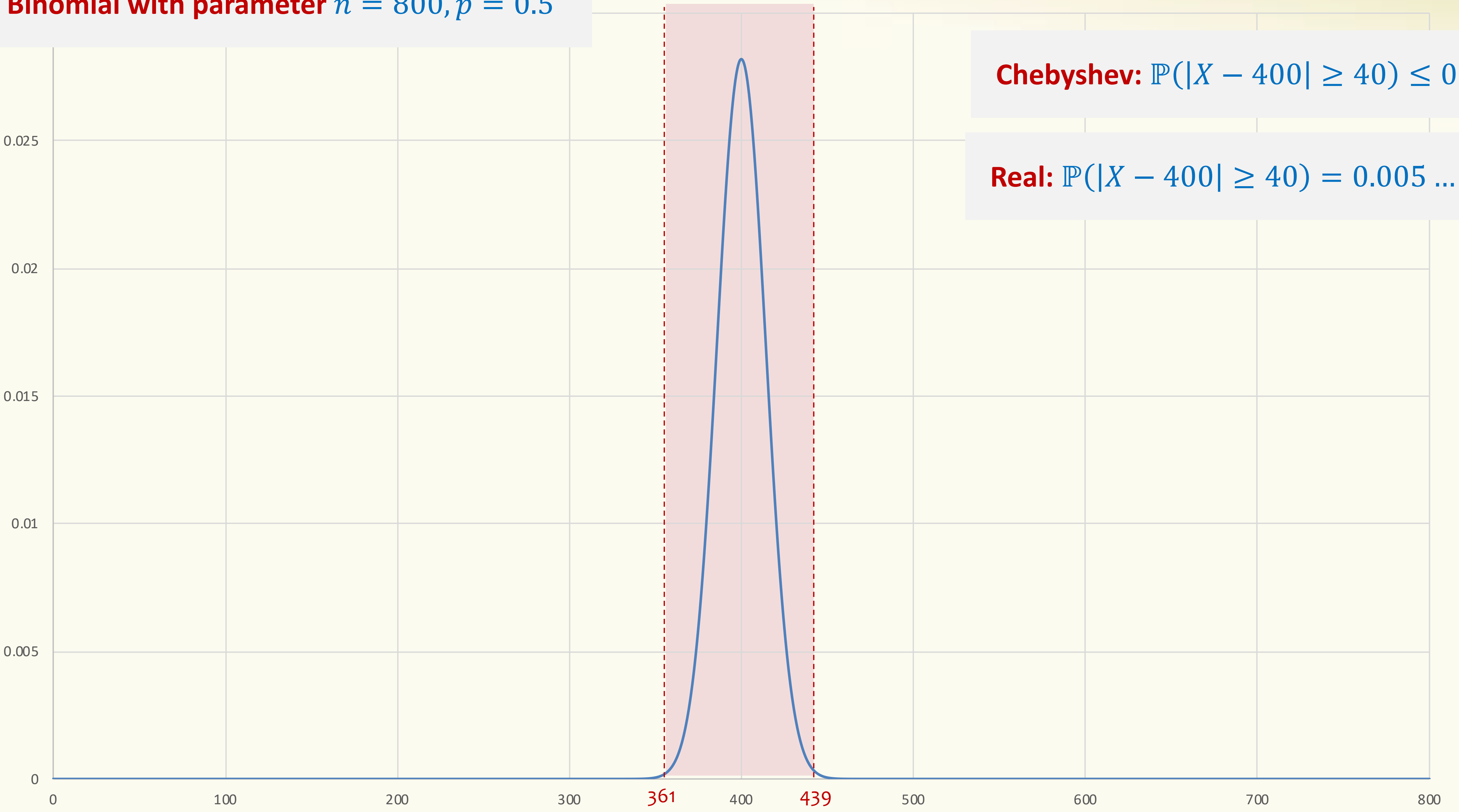
Binomial with parameter $n = 200, p = 0.5$

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

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



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



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

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

Can we do better?

- Chebyshev's inequality indicates that the probability that we are off by at least $\epsilon \cdot \mathbb{E}(X)$ goes to 0 as $\frac{1-p}{\epsilon^2 np} = O(1/n)$ for fixed p and ϵ
- Exact analysis indicates that probability goes to 0 much faster, at least for a binomial random variable.
  – How fast?

Chernoff-Hoeffding Bound – Binomial Distribution

Theorem. (Chernoff-Hoeffding) Let X be a binomial RV with parameters p and n. Let $\mu = np = \mathbb{E}(X)$. Then, for any $\delta > 0$,

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

Binomial: $n = 800, p = 0.5 \rightarrow \mu = np = 400$

Chebyshev: $\mathbb{P}(|X - \mu| \geq 0.1\mu) \leq 0.125$

CH: $\mathbb{P}(|X - \mu| \geq 0.1\mu) \leq 2e^{-\frac{4}{2.1}} = 0.296 \ldots$

Chernoff-Hoeffding Bound – Binomial Distribution

Theorem. (Chernoff-Hoeffding) Let X be a binomial RV with parameters p and n. Let $\mu = np = \mathbb{E}(X)$. Then, for any $\delta > 0$,

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

Binomial: $n = 8000, p = 0.5 \rightarrow \mu = np = 4000$

Chebyshev: $\mathbb{P}(|X - \mu| \geq 0.1\mu) \leq 0.0125$

CH: $\mathbb{P}(|X - \mu| \geq 0.1\mu) \leq 2e^{-\frac{40}{2.1}} \approx 1.7\times 10^{-8}$