DCP 1206: Probability
Lecture 22 — Chernoff Bounds, Hoeffding's Inequality, and WLLN

Ping-Chun Hsieh

December 6, 2019

Announcements

- Our class will be relocated at <u>EC114</u> next Wednesday (12/11)

- HW5 is posted on E3 (Due: 12/11 in class)

This Lecture

1. Chernoff Bounds and Hoeffding's Inequality

2. Weak Law of Large Numbers

- Reading material: Chapter 11.1-11.3

Review: Chebyshev's Inequality

- **Chebyshev's Inequality**: Let X be a random variable with mean μ and variance σ^2. Then, for any $t > 0$,

$$P(|X - \mu| \geq t) \leq \frac{\sigma^2}{t^2}$$

- Proof:

Review: Chebyshev's Inequality and Sample Mean

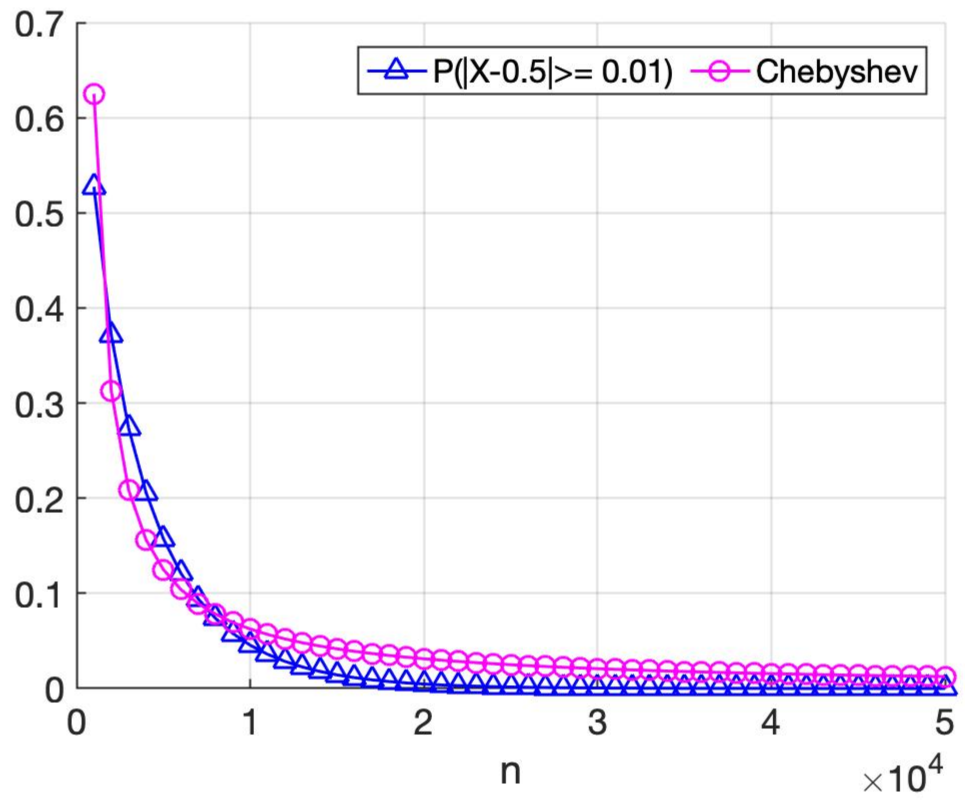
- **Chebyshev's and Sample Mean**: Let $X_1, \cdots, X_n$ be a sequence of independent and identically distributed (i.i.d.) random variables with mean μ and variance σ^2. Define $\bar{X} = \frac{1}{n}(X_1 + \cdots + X_n)$. Then, for any $\varepsilon > 0$, we have

$$P(|\bar{X} - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2 n}$$

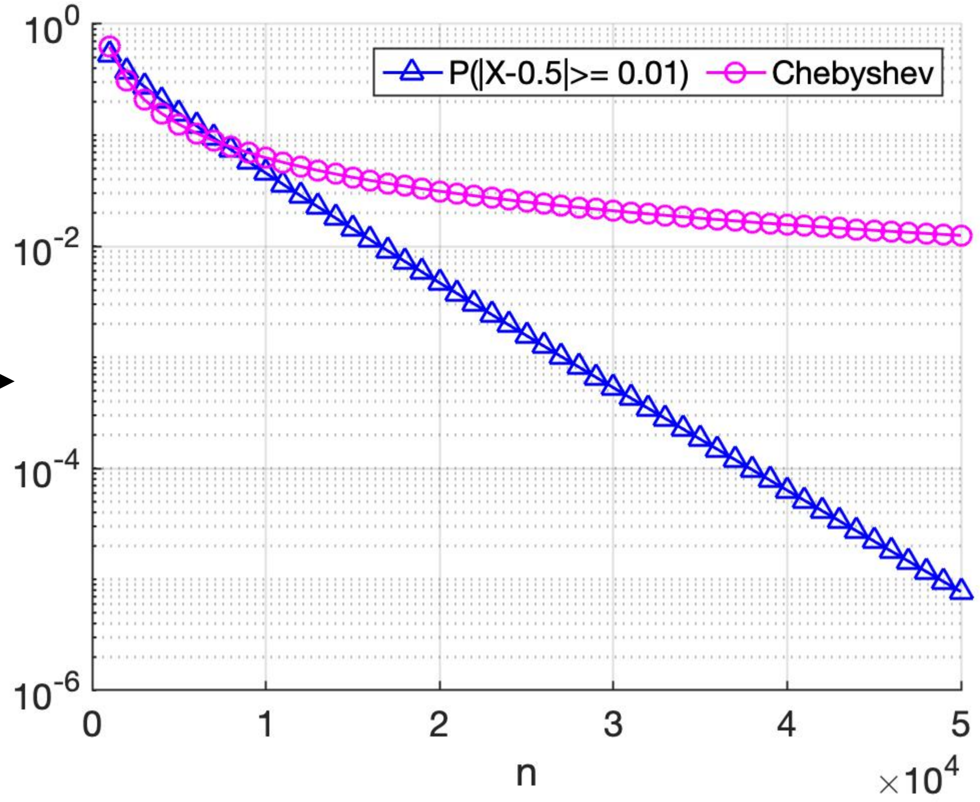
Any Issue With Chebyshev's Inequality?

- Example: $X_1, \cdots, X_n$ are i.i.d. Bernoulli with parameter 0.5
  - $E[X_i] =$ ________ and Var$[X_i] =$ _________
  - Chebyshev's: $P(|\bar{X} - \mu| \geq \varepsilon) \leq \dfrac{\sigma^2}{\varepsilon^2 n}$
  - Let's plot $P(|\bar{X} - \mu| \geq \varepsilon)$ for small ε

$\underline{\varepsilon = 0.01}$



log scale →



1. Chernoff Bounds

Chernoff Bound

- **Chernoff Bound**: Let X be a random variable with MGF $M_X(t)$ Suppose $M_X(t)$ exists for all t in some set S. Then, for any $t > 0$ and $t \in S$, for any $a \in \mathbb{R}$, we have

$$P(X \geq a) \leq e^{-ta} \cdot M_X(t)$$

- Proof:

Optimizing the Chernoff Bound

- **Chernoff Bound**: Let X be a random variable with MGF $M_X(t)$ Suppose $M_X(t)$ exists for all t in some set S. Then, for any $t > 0$ and $t \in S$, for any $a \in \mathbb{R}$, we have

$$P(X \geq a) \leq e^{-\phi(a)},$$

where $\phi(a) = \max_{t>0, t\in S} (ta - \ln M_X(t))$

- Proof:

Example: Chernoff Bound for Bernoulli R.V.s

- Example: Suppose $X \sim$ Bernoulli(p)
  - What is $M_X(t)$?
  - What is the Chernoff bound for X? ($P(X \geq a) \leq e^{-ta} \cdot M_X(t)$)

Example: Optimizing Chernoff Bound for Bernoulli R.V.s

- Example: Suppose $X \sim$ Bernoulli(p)
  - How to optimize the Chernoff bound for X?
    $(P(X \geq a) \leq e^{-\phi(a)}, \phi(a) = \max_{t>0, t\in S} (ta - \ln M_X(t)))$

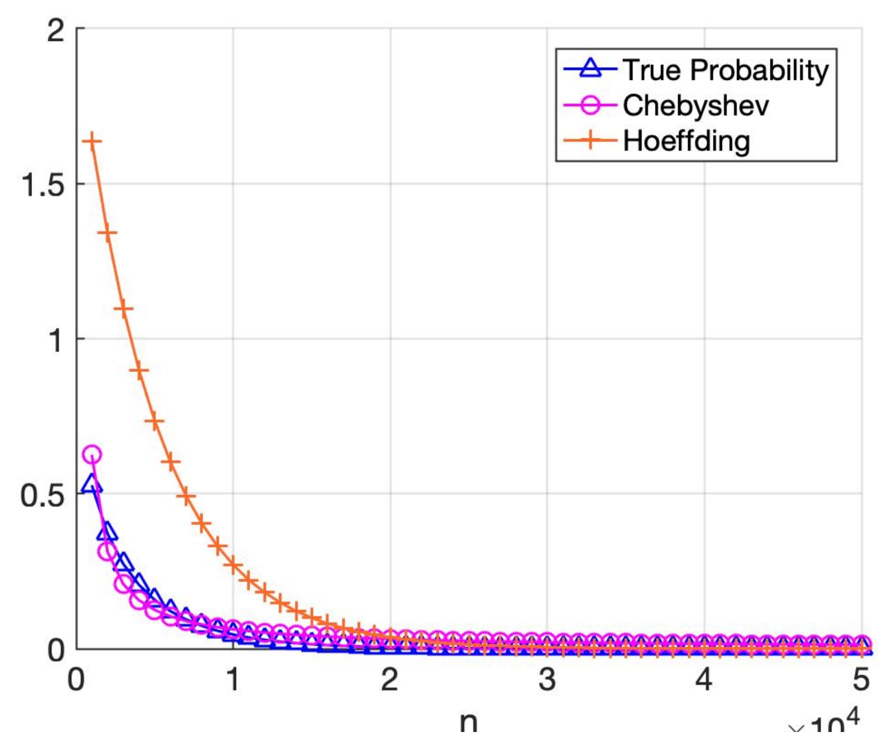
How about applying Chernoff bound to sum of independent random variables?

Hoeffding's Inequality (Formally)

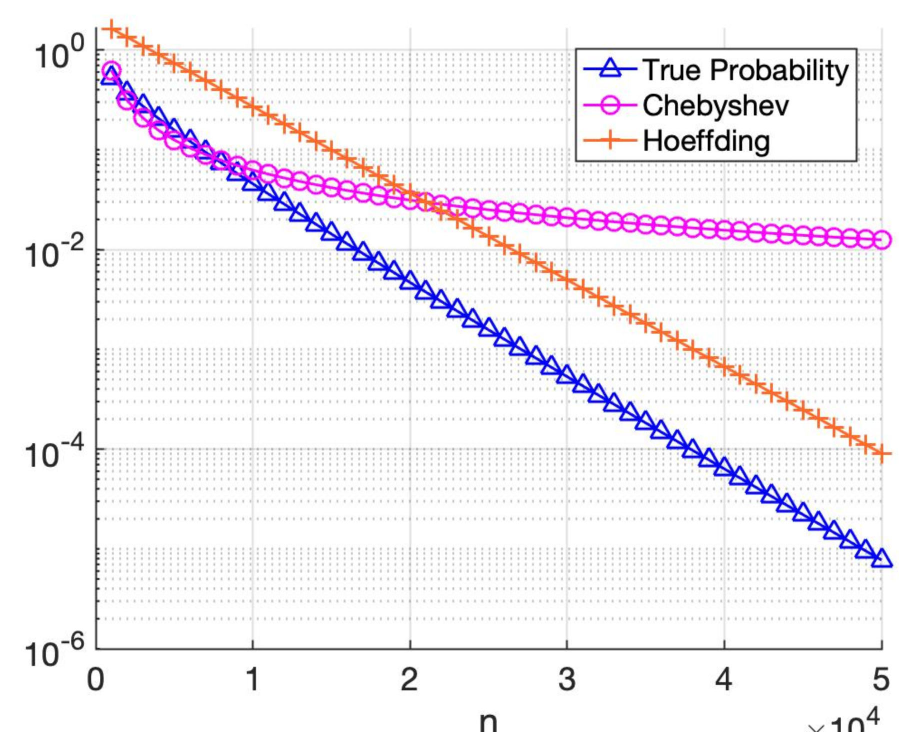
- **Hoeffding's Inequality (For Bernoulli)**: Let $X_1, \cdots, X_n$ be a sequence of i.i.d. Bernoulli random variables with parameter p. Define $\bar{X} = \frac{1}{n}(X_1 + \cdots + X_n)$. Then, for any $\varepsilon > 0$, we have

$$P(|\bar{X} - p| \geq \varepsilon) \leq 2\exp(-2n\varepsilon^2)$$

$\underline{\varepsilon = 0.01}$



log scale →



Proof of Hoeffding's Inequality (Positive Part)

$$P(\bar{X} - p \geq \varepsilon) \leq \exp(-2n\varepsilon^2)$$

- [Hint] Chernoff bound: $P(X \geq a) \leq e^{-ta} \cdot M_X(t)$

$$P(\bar{X} - p \geq \varepsilon) \leq$$

Hoeffding's Lemma

- **Hoeffding's Lemma**: Let Z be a random variable with $E[Z] = 0$, and $Z \in [a, b]$ with probability 1. Then, for any $t > 0$, we have
$$E[e^{tZ}] \leq \exp\left(\frac{t^2(b-a)^2}{8}\right)$$

- Question: If $Z \sim \text{Bernoulli}(p)$, then $E[e^{t(Z-p)}] \leq$

Proof of Hoeffding's Inequality (Negative Part)

$$P(\bar{X} - p \leq -\varepsilon) = P(p - \bar{X} \geq \varepsilon) \leq \exp(-2n\varepsilon^2)$$

- [Hint] Chernoff bound: $P(X \geq a) \leq e^{-ta} \cdot M_X(t)$

$$P(p - \bar{X} \geq \varepsilon) \leq$$

What if we let $n \to \infty$?

Chebyshev's and Sample Mean: $n \to \infty$

- **Chebyshev's and Sample Mean**: Let $X_1, \cdots, X_n$ be a sequence of independent and identically distributed (i.i.d.) random variables with mean μ and variance σ^2. Define $\bar{X} = \frac{1}{n}(X_1 + \cdots + X_n)$. Then, for any $\varepsilon > 0$, we have

$$P(|\bar{X} - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2 n}$$

- By letting $n \to \infty$:

The Weak Law of Large Numbers (WLLN)

- **The Weak Law of Large Numbers (Khinchin's Law)**: Let $X_1, \cdots, X_n$ be a sequence of independent and identically distributed (i.i.d.) random variables with mean μ. Define $\bar{X} = \frac{1}{n}(X_1 + \cdots + X_n)$. Then, for every $\varepsilon > 0$, we have

$$P(|\bar{X} - \mu| \geq \varepsilon) \to 0 \qquad \text{as } n \to \infty$$

- Question: Any change in technical conditions (cf: Chebyshev's)?

- Question: What does "convergence" mean here?

Convergence in Probability

- **Convergence of a Deterministic Sequence**: Let $a_1, a_2 \cdots$ be a sequence of real numbers. We say that a_n converges to a if for every $\varepsilon > 0$, there exists N_0 such that

$$|a_n - a| \leq \varepsilon \qquad \text{for all } n \geq N_0$$

- **Convergence in Probability**: Let $Y_1, Y_2 \cdots$ be a sequence of i.i.d. random variables, and let a be a real number. We say that Y_n converges to a if for every $\varepsilon > 0$, we have

Next Lecture

- Strong Law of Large Numbers
- Central Limit Theorem

1-Minute Summary

1. Chernoff Bounds and Hoeffding's Inequality

- Put the r.v. in the exponentiate (similar to MGF)
- Hoeffding's: $P(|\bar{X} - p| \geq \varepsilon) \leq 2\exp(-2n\varepsilon^2)$

2. Weak Law of Large Numbers

- Convergence in probability