WLLN, Confidence Intervals (CI): Chebyshev vs. CLT
WLLN, Confidence Intervals (CI): Chebyshev vs. CLT
1. Review: Inequalities: Markov, Chebyshev
2. Law of Large Numbers
3. Review: CLT
4. Confidence Intervals: Chebyshev vs. CLT
Inequalities: An Overview

**Distribution**

**Markov**

**Chebyshev**

\[ Pr\left[ X > a \right] \]

\[ Pr\left[ |X - \mu| > \epsilon \right] \]
Markov Inequality

If $X$ can only take non-negative values then

$$P(X \geq a) \leq \frac{E[X]}{a}$$

for all $a > 0$.

This inequality makes no assumptions on the existence of variance and so it can’t be very strong for typical distributions. In fact, it is quite weak.
Chebyshev Inequality

If $X$ is a random variable with finite mean and variance $\sigma^2$, then

$$P(|X - E[X]| \geq c) \leq \frac{\sigma^2}{c^2}$$

for all $c > 0$.

Also, letting $c = k\sigma$:

$$P(|X - E[X]| \geq k\sigma) \leq \frac{1}{k^2}$$
Fraction of $H$’s

Here is a classical application of Chebyshev’s inequality.

Let $X_m = 1$ if the $m$-th flip of a fair coin is $H$ and $X_m = 0$ otherwise.

Define $M_n = X_1 + \cdots + X_n$, for $n \geq 1$.

We want to estimate $\Pr[|M_n - 0.5| \geq 0.1] = \Pr[M_n \leq 0.4 \text{ or } M_n \geq 0.6]$.

By Chebyshev, $\Pr[|M_n - 0.5| \geq 0.1] \leq \frac{\text{var}[M_n]}{(0.1)^2} = 100 \times \frac{\text{var}[X_1]}{n}$.

Now, $\text{var}[M_n] = \frac{1}{n} \sum_{i=1}^{n} \text{var}[X_i] \leq \frac{0.5 \times 0.5}{4n}$. 

$\text{Var}(X_i) = p(1-p) \leq (0.5)(0.5) = \frac{1}{4}$.
Fraction of $H$’s

Here is a classical application of Chebyshev’s inequality. How likely is it that the fraction of $H$’s differs from 50%?
Here is a classical application of Chebyshev’s inequality. How likely is it that the fraction of $H$’s differs from 50%?

Let $X_m = 1$ if the $m$-th flip of a fair coin is $H$ and $X_m = 0$ otherwise.
Fraction of *H*'s

Here is a classical application of Chebyshev’s inequality. How likely is it that the fraction of *H*'s differs from 50%?

Let $X_m = 1$ if the $m$-th flip of a fair coin is $H$ and $X_m = 0$ otherwise.

Define

$$M_n = \frac{X_1 + \cdots + X_n}{n}, \text{ for } n \geq 1.$$
Fraction of $H$’s

Here is a classical application of Chebyshev’s inequality. How likely is it that the fraction of $H$’s differs from 50%?

Let $X_m = 1$ if the $m$-th flip of a fair coin is $H$ and $X_m = 0$ otherwise.

Define

$$M_n = \frac{X_1 + \cdots + X_n}{n}, \text{ for } n \geq 1.$$ 

We want to estimate

$$Pr[|M_n - 0.5| \geq 0.1] = Pr[M_n \leq 0.4 \text{ or } M_n \geq 0.6].$$
Fraction of $H$’s

Here is a classical application of Chebyshev’s inequality.

How likely is it that the fraction of $H$’s differs from 50%?

Let $X_m = 1$ if the $m$-th flip of a fair coin is $H$ and $X_m = 0$ otherwise.

Define

$$M_n = \frac{X_1 + \cdots + X_n}{n}, \text{ for } n \geq 1.$$

We want to estimate

$$Pr[|M_n - 0.5| \geq 0.1] = Pr[M_n \leq 0.4 \text{ or } M_n \geq 0.6].$$

By Chebyshev,
Fraction of $H$’s

Here is a classical application of Chebyshev’s inequality. How likely is it that the fraction of $H$’s differs from 50%?

Let $X_m = 1$ if the $m$-th flip of a fair coin is $H$ and $X_m = 0$ otherwise.

Define

$$M_n = \frac{X_1 + \cdots + X_n}{n}, \text{ for } n \geq 1.$$ 

We want to estimate

$$Pr[|M_n - 0.5| \geq 0.1] = Pr[M_n \leq 0.4 \text{ or } M_n \geq 0.6].$$

By Chebyshev,

$$Pr[|M_n - 0.5| \geq 0.1] \leq \frac{\text{var}[M_n]}{(0.1)^2} = 100\text{var}[M_n].$$
Fraction of $H$’s

Here is a classical application of Chebyshev’s inequality.

How likely is it that the fraction of $H$’s differs from 50%?

Let $X_m = 1$ if the $m$-th flip of a fair coin is $H$ and $X_m = 0$ otherwise.

Define

$$M_n = \frac{X_1 + \cdots + X_n}{n}, \text{ for } n \geq 1.$$  

We want to estimate

$$Pr[|M_n - 0.5| \geq 0.1] = Pr[M_n \leq 0.4 \text{ or } M_n \geq 0.6].$$

By Chebyshev,

$$Pr[|M_n - 0.5| \geq 0.1] \leq \frac{\text{var}[M_n]}{(0.1)^2} = 100\text{var}[M_n].$$

Now,
Here is a classical application of Chebyshev’s inequality.

How likely is it that the fraction of $H$’s differs from 50%?

Let $X_m = 1$ if the $m$-th flip of a fair coin is $H$ and $X_m = 0$ otherwise.

Define

$$M_n = \frac{X_1 + \cdots + X_n}{n}, \text{ for } n \geq 1.$$ 

We want to estimate

$$Pr[|M_n - 0.5| \geq 0.1] = Pr[M_n \leq 0.4 \text{ or } M_n \geq 0.6].$$

By Chebyshev,

$$Pr[|M_n - 0.5| \geq 0.1] \leq \frac{\text{var}[M_n]}{(0.1)^2} = 100\text{var}[M_n].$$

Now,

$$\text{var}[M_n] = \frac{1}{n^2}(\text{var}[X_1] + \cdots + \text{var}[X_n])$$
Fraction of $H$'s

Here is a classical application of Chebyshev’s inequality. How likely is it that the fraction of $H$’s differs from 50%?

Let $X_m = 1$ if the $m$-th flip of a fair coin is $H$ and $X_m = 0$ otherwise.

Define

$$M_n = \frac{X_1 + \cdots + X_n}{n}, \text{ for } n \geq 1.$$ 

We want to estimate

$$Pr[|M_n - 0.5| \geq 0.1] = Pr[M_n \leq 0.4 \text{ or } M_n \geq 0.6].$$

By Chebyshev,

$$Pr[|M_n - 0.5| \geq 0.1] \leq \frac{\text{var}[M_n]}{(0.1)^2} = 100\text{var}[M_n].$$

Now,

$$\text{var}[M_n] = \frac{1}{n^2} (\text{var}[X_1] + \cdots + \text{var}[X_n]) = \frac{1}{n} \text{var}[X_1]$$
Fraction of $H$’s

Here is a classical application of Chebyshev’s inequality.

How likely is it that the fraction of $H$’s differs from 50%?

Let $X_m = 1$ if the $m$-th flip of a fair coin is $H$ and $X_m = 0$ otherwise.

Define

$$M_n = \frac{X_1 + \cdots + X_n}{n}, \text{ for } n \geq 1.$$ 

We want to estimate

$$Pr[|M_n - 0.5| \geq 0.1] = Pr[M_n \leq 0.4 \text{ or } M_n \geq 0.6].$$

By Chebyshev,

$$Pr[|M_n - 0.5| \geq 0.1] \leq \frac{\text{var}[M_n]}{(0.1)^2} = 100 \text{var}[M_n].$$

Now,

$$\text{var}[M_n] = \frac{1}{n^2} (\text{var}[X_1] + \cdots + \text{var}[X_n]) = \frac{1}{n} \text{var}[X_1] \leq \frac{1}{4n}.$$
Fraction of $H$’s

Here is a classical application of Chebyshev’s inequality. How likely is it that the fraction of $H$’s differs from 50%?

Let $X_m = 1$ if the $m$-th flip of a fair coin is $H$ and $X_m = 0$ otherwise.

Define

$$M_n = \frac{X_1 + \cdots + X_n}{n}, \text{ for } n \geq 1.$$ 

We want to estimate

$$Pr[|M_n - 0.5| \geq 0.1] = Pr[M_n \leq 0.4 \text{ or } M_n \geq 0.6].$$

By Chebyshev,

$$Pr[|M_n - 0.5| \geq 0.1] \leq \frac{\text{var}[M_n]}{(0.1)^2} = 100\text{var}[M_n].$$

Now,

$$\text{var}[M_n] = \frac{1}{n^2}(\text{var}[X_1] + \cdots + \text{var}[X_n]) = \frac{1}{n}\text{var}[X_1] \leq \frac{1}{4n}.$$ 

$$\text{Var}(X_i) = p(1 - p) \leq (.5)(.5) = \frac{1}{4}$$
Fraction of $H$’s

\[ M_n = \frac{X_1 + \cdots + X_n}{n}, \quad \text{for } n \geq 1. \]

\[ \Pr[|M_n - 0.5| \geq 0.1] \leq \frac{25}{n}. \]
Fraction of $H$’s

\[ M_n = \frac{X_1 + \cdots + X_n}{n}, \text{ for } n \geq 1. \]

\[ Pr[|M_n - 0.5| \geq 0.1] \leq \frac{25}{n}. \]

For $n = 1,000$, we find that this probability is less than 2.5%.
Fraction of $H$'s

\[ M_n = \frac{X_1 + \cdots + X_n}{n}, \text{ for } n \geq 1. \]

\[ Pr[|M_n - 0.5| \geq 0.1] \leq \frac{25}{n}. \]

For $n = 1,000$, we find that this probability is less than 2.5%.

As $n \to \infty$, this probability goes to zero.
Fraction of $H$’s

$$M_n = \frac{X_1 + \cdots + X_n}{n}, \text{ for } n \geq 1.$$ 

$$Pr[|M_n - 0.5| \geq 0.1] \leq \frac{25}{n}.$$ 

For $n = 1,000$, we find that this probability is less than 2.5%. As $n \to \infty$, this probability goes to zero.

In fact, for any $\varepsilon > 0$, 

...
Fraction of $H$'s

$$M_n = \frac{X_1 + \cdots + X_n}{n}, \text{ for } n \geq 1.$$  

$$Pr[|M_n - 0.5| \geq 0.1] \leq \frac{25}{n}.$$  

For $n = 1,000$, we find that this probability is less than 2.5%.  
As $n \to \infty$, this probability goes to zero.  
In fact, for any $\varepsilon > 0$, as $n \to \infty$, 
Fraction of $H$’s

$$M_n = \frac{X_1 + \cdots + X_n}{n}, \text{ for } n \geq 1.$$  

$$Pr[|M_n - 0.5| \geq 0.1] \leq \frac{25}{n}.$$  

For $n = 1,000$, we find that this probability is less than 2.5%. As $n \to \infty$, this probability goes to zero.

In fact, for any $\varepsilon > 0$, as $n \to \infty$, the probability that the fraction of $H$s is within $\varepsilon > 0$ of 50% approaches 1:
Fraction of $H$’s

\[ M_n = \frac{X_1 + \cdots + X_n}{n}, \text{ for } n \geq 1. \]

\[ Pr[|M_n - 0.5| \geq 0.1] \leq \frac{25}{n}. \]

For $n = 1,000$, we find that this probability is less than 2.5%.

As $n \to \infty$, this probability goes to zero.

In fact, for any $\varepsilon > 0$, as $n \to \infty$, the probability that the fraction of $H$s is within $\varepsilon > 0$ of 50% approaches 1:

\[ Pr[|M_n - 0.5| \leq \varepsilon] \to 1. \]
Fraction of H’s

\[ M_n = \frac{X_1 + \cdots + X_n}{n}, \text{ for } n \geq 1. \]

\[ Pr[|M_n - 0.5| \geq 0.1] \leq \frac{25}{n}. \]

For \( n = 1,000 \), we find that this probability is less than 2.5%.

As \( n \to \infty \), this probability goes to zero.

In fact, for any \( \varepsilon > 0 \), as \( n \to \infty \), the probability that the fraction of Hs is within \( \varepsilon > 0 \) of 50% approaches 1:

\[ Pr[|M_n - 0.5| \leq \varepsilon] \to 1. \]

This is an example of the (Weak) Law of Large Numbers.
Fraction of $H$’s

\[ M_n = \frac{X_1 + \cdots + X_n}{n}, \text{ for } n \geq 1. \]

\[ Pr[|M_n - 0.5| \geq 0.1] \leq \frac{25}{n}. \]

For $n = 1,000$, we find that this probability is less than 2.5%. As $n \to \infty$, this probability goes to zero.

In fact, for any $\varepsilon > 0$, as $n \to \infty$, the probability that the fraction of $H$s is within $\varepsilon > 0$ of 50% approaches 1:

\[ Pr[|M_n - 0.5| \leq \varepsilon] \to 1. \]

This is an example of the (Weak) Law of Large Numbers.

We look at a general case next.
We perform an experiment $n$ times independently and

$$M_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

The fact that $\text{var}(M_n) \to 0$ at rate $\frac{1}{n}$ is great but what does that tell us about $P(|M_n - E[X_i]|)$? How quickly does it go to zero? Just use Chebyshev:

$$P(|X - E[X]| \geq c) \leq \frac{\sigma^2}{c^2}$$

$$P(|M_n - E[X_i]| \geq \epsilon) \leq \frac{\sigma^2}{n\epsilon^2}$$

for any $\epsilon > 0$.

This is a form of the Weak Law of Large Numbers.
Weak Law of Large Numbers

**Theorem** Weak Law of Large Numbers
Weak Law of Large Numbers

**Theorem** Weak Law of Large Numbers

Let $X_1, X_2, \ldots$ be pairwise independent with the same distribution and mean $\mu$. 
Weak Law of Large Numbers

**Theorem** Weak Law of Large Numbers

Let $X_1, X_2, \ldots$ be pairwise independent with the same distribution and mean $\mu$. Then, for all $\varepsilon > 0$,

$$Pr[\left| \frac{X_1 + \cdots + X_n}{n} - \mu \right| \geq \varepsilon] \rightarrow 0, \text{ as } n \rightarrow \infty.$$
Weak Law of Large Numbers

**Theorem** Weak Law of Large Numbers

Let \( X_1, X_2, \ldots \) be pairwise independent with the same distribution and mean \( \mu \). Then, for all \( \varepsilon > 0 \),

\[
Pr\left[ \left| \frac{X_1 + \cdots + X_n}{n} - \mu \right| \geq \varepsilon \right] \to 0, \quad \text{as } n \to \infty.
\]

**Proof:**

Let \( M_n = \frac{X_1 + \cdots + X_n}{n} \). Then
Theorem Weak Law of Large Numbers

Let $X_1, X_2, \ldots$ be pairwise independent with the same distribution and mean $\mu$. Then, for all $\varepsilon > 0$,

$$Pr[|\frac{X_1 + \cdots + X_n}{n} - \mu| \geq \varepsilon] \to 0, \text{ as } n \to \infty.$$  

Proof:

Let $M_n = \frac{X_1 + \cdots + X_n}{n}$. Then

$$Pr[|M_n - \mu| \geq \varepsilon] \leq \frac{\text{var}[M_n]}{\varepsilon^2}$$
Weak Law of Large Numbers

**Theorem** Weak Law of Large Numbers

Let $X_1, X_2, \ldots$ be pairwise independent with the same distribution and mean $\mu$. Then, for all $\varepsilon > 0$,

$$\Pr\left[\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| \geq \varepsilon\right] \to 0, \text{ as } n \to \infty.$$ 

**Proof:**

Let $M_n = \frac{X_1 + \cdots + X_n}{n}$. Then

$$\Pr[|M_n - \mu| \geq \varepsilon] \leq \frac{\text{var}[M_n]}{\varepsilon^2} = \frac{\text{var}[X_1 + \cdots + X_n]}{n^2 \varepsilon^2}$$
**Theorem** Weak Law of Large Numbers

Let $X_1, X_2, \ldots$ be pairwise independent with the same distribution and mean $\mu$. Then, for all $\varepsilon > 0$,

$$\Pr[|\frac{X_1 + \cdots + X_n}{n} - \mu| \geq \varepsilon] \to 0, \text{ as } n \to \infty.$$  

**Proof:**
Let $M_n = \frac{X_1 + \cdots + X_n}{n}$. Then

$$\Pr[|M_n - \mu| \geq \varepsilon] \leq \frac{\text{var}[M_n]}{\varepsilon^2} = \frac{\text{var}[X_1 + \cdots + X_n]}{n^2 \varepsilon^2} = \frac{n \text{var}[X_1]}{n^2 \varepsilon^2}$$
Weak Law of Large Numbers

**Theorem** Weak Law of Large Numbers

Let $X_1, X_2, \ldots$ be pairwise independent with the same distribution and mean $\mu$. Then, for all $\varepsilon > 0$,

$$Pr\left[\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| \geq \varepsilon \right] \to 0, \text{ as } n \to \infty.$$

**Proof:**

Let $M_n = \frac{X_1 + \cdots + X_n}{n}$. Then

$$Pr\left[|M_n - \mu| \geq \varepsilon \right] \leq \frac{\text{var}[M_n]}{\varepsilon^2} = \frac{\text{var}[X_1 + \cdots + X_n]}{n^2 \varepsilon^2} = \frac{n \text{var}[X_1]}{n^2 \varepsilon^2} = \frac{\text{var}[X_1]}{n \varepsilon^2}.$$
**Weak Law of Large Numbers**

**Theorem** Weak Law of Large Numbers

Let $X_1, X_2, \ldots$ be pairwise independent with the same distribution and mean $\mu$. Then, for all $\varepsilon > 0$,

$$Pr\left[\left| \frac{X_1 + \cdots + X_n}{n} - \mu \right| \geq \varepsilon \right] \rightarrow 0, \text{ as } n \rightarrow \infty.$$  

**Proof:**

Let $M_n = \frac{X_1 + \cdots + X_n}{n}$. Then

$$Pr[|M_n - \mu| \geq \varepsilon] \leq \frac{\text{var}[M_n]}{\varepsilon^2} = \frac{\text{var}[X_1 + \cdots + X_n]}{n^2 \varepsilon^2}$$

$$= \frac{n \text{var}[X_1]}{n^2 \varepsilon^2} = \frac{\text{var}[X_1]}{n \varepsilon^2} \rightarrow 0, \text{ as } n \rightarrow \infty.$$
What does the Weak Law Really Mean?

WLLN: \( \lim_{n \to \infty} P(|M_n - \mu| \geq \epsilon) = 0. \)

Just using the defn of limit: For any \( \epsilon, \delta > 0 \), there exists a number \( n(\epsilon, \delta) \) such that

\[
P(|M_n - \mu| \geq \epsilon) \leq \delta \quad \text{for all } n \geq n(\epsilon, \delta)
\]

- \( \delta \): Confidence level
- \( \epsilon \): "Error"
- \( n(\epsilon, \delta) \): threshold function for a given level of confidence and accuracy

What this is saying is that if we compute \( M_n \) for large \( n \) then:

Almost Always, \( |M_n - \mu| < \epsilon. \)

We say that \( M_n \) converges to \( \mu \) in probability.
Recap: Normal (Gaussian) Distribution.

For any $\mu$ and $\sigma$, a normal (aka Gaussian) random variable $Y$, which we write as $Y = \mathcal{N}(\mu, \sigma^2)$, has pdf

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-\mu)^2 / 2\sigma^2}.$$ 

Standard normal has $\mu = 0$ and $\sigma = 1$.

Note: $Pr[|Y - \mu| > 1.65\sigma] = 10\%$; $Pr[|Y - \mu| > 2\sigma] = 5\%$. 

![Normal distribution graph with annotations for standard deviations and probabilities.](image-url)
Recap: Central Limit Theorem

Let $X_1, X_2, \ldots$ be i.i.d. with $E[X_1] = \mu$ and $\text{var}(X_1) = \sigma^2$. Define $S_n := \frac{A_n - \mu}{\sigma \sqrt{n}} = X_1 + \cdots + X_n - n \mu / \sigma \sqrt{n}$.

Then, $S_n \to N(0, 1)$, as $n \to \infty$.

That is, $\Pr[S_n \leq \alpha] \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-x^2/2} dx$.

$E(S_n) = 1/\sigma \sqrt{n}(E(A_n) - \mu) = 0$.

$\text{Var}(S_n) = 1/\sigma^2 / n \text{Var}(A_n) = 1/\sigma^2$.
Recap: Central Limit Theorem

Central Limit Theorem
Let \( X_1, X_2, \ldots \) be i.i.d. with \( E[X_1] = \mu \) and \( \text{var}(X_1) = \sigma^2 \). Define

\[
S_n := \frac{A_n - \mu}{\sigma / \sqrt{n}} = \frac{X_1 + \cdots + X_n - n\mu}{\sigma \sqrt{n}}.
\]

Then,

\[
S_n \to \mathcal{N}(0, 1), \text{ as } n \to \infty.
\]

That is,

\[
E(S_n) = \frac{1}{\sigma / \sqrt{n}}(E(A_n) - \mu) = 0
\]

\[
\text{Var}(S_n) = \frac{1}{\sigma^2 / n} \text{Var}(A_n) = 1.
\]

\[
Pr[S_n \leq \alpha] \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-x^2/2} \, dx.
\]
Confidence Interval (CI) for Mean: CLT

Let $X_1, X_2, \ldots$ be i.i.d. with mean $\mu$ and variance $\sigma^2$.

Let $A_n = X_1 + \cdots + X_n$.

The CLT states that $A_n - \mu \sigma / \sqrt{n} \rightarrow N(0, 1)$ as $n \rightarrow \infty$.

Thus, for $n \gg 1$, one has $\Pr[-2 \leq (A_n - \mu \sigma / \sqrt{n}) \leq 2] \approx 95\%$.

Equivalently, $\Pr[\mu \in [A_n - 2 \sigma / \sqrt{n}, A_n + 2 \sigma / \sqrt{n}]] \approx 95\%$.

That is, $[A_n - 2 \sigma / \sqrt{n}, A_n + 2 \sigma / \sqrt{n}]$ is a 95\% CI for $\mu$. 
Confidence Interval (CI) for Mean: CLT

Let $X_1, X_2, \ldots$ be i.i.d. with mean $\mu$ and variance $\sigma^2$. 
Confidence Interval (CI) for Mean: CLT

Let \( X_1, X_2, \ldots \) be i.i.d. with mean \( \mu \) and variance \( \sigma^2 \). Let

\[
A_n = \frac{X_1 + \cdots + X_n}{n}.
\]
Confidence Interval (CI) for Mean: CLT

Let $X_1, X_2, \ldots$ be i.i.d. with mean $\mu$ and variance $\sigma^2$. Let

$$A_n = \frac{X_1 + \cdots + X_n}{n}.$$ 

The CLT states that

$$\frac{A_n - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \rightarrow \mathcal{N}(0,1) \text{ as } n \rightarrow \infty.$$
Confidence Interval (CI) for Mean: CLT

Let $X_1, X_2, \ldots$ be i.i.d. with mean $\mu$ and variance $\sigma^2$. Let

$$A_n = \frac{X_1 + \cdots + X_n}{n}.$$ 

The CLT states that

$$\frac{A_n - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \rightarrow \mathcal{N}(0, 1) \text{ as } n \rightarrow \infty.$$ 

Thus, for $n \gg 1$, one has

$$Pr[-2 \leq \left(\frac{A_n - \mu}{\sigma/\sqrt{n}}\right) \leq 2] \approx 95\%.$$
Confidence Interval (CI) for Mean: CLT

Let $X_1, X_2, \ldots$ be i.i.d. with mean $\mu$ and variance $\sigma^2$. Let

$$A_n = \frac{X_1 + \cdots + X_n}{n}.$$ 

The CLT states that

$$\frac{A_n - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \to \mathcal{N}(0,1) \text{ as } n \to \infty.$$ 

Thus, for $n \gg 1$, one has

$$\Pr[-2 \leq \left( \frac{A_n - \mu}{\sigma/\sqrt{n}} \right) \leq 2] \approx 95\%.$$ 

Equivalently,

$$\Pr[\mu \in [A_n - 2\frac{\sigma}{\sqrt{n}}, A_n + 2\frac{\sigma}{\sqrt{n}}]] \approx 95\%.$$
Confidence Interval (CI) for Mean: CLT

Let $X_1, X_2, \ldots$ be i.i.d. with mean $\mu$ and variance $\sigma^2$. Let

$$A_n = \frac{X_1 + \cdots + X_n}{n}.$$ 

The CLT states that

$$\frac{A_n - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \to \mathcal{N}(0, 1) \text{ as } n \to \infty.$$ 

Thus, for $n \gg 1$, one has

$$Pr[-2 \leq \left( \frac{A_n - \mu}{\sigma/\sqrt{n}} \right) \leq 2] \approx 95\%.$$ 

Equivalently,

$$Pr[\mu \in [A_n - 2 \frac{\sigma}{\sqrt{n}}, A_n + 2 \frac{\sigma}{\sqrt{n}}]] \approx 95\%.$$ 

That is,

$$[A_n - 2 \frac{\sigma}{\sqrt{n}}, A_n + 2 \frac{\sigma}{\sqrt{n}}] \text{ is a } 95\% \text{ - CI for } \mu.$$
CI for Mean: CLT vs. Chebyshev

Let $X_1, X_2, \ldots$ be i.i.d. with mean $\mu$ and variance $\sigma^2$. Let

$$A_n = \frac{X_1 + \cdots + X_n}{n}.$$ 

The CLT states that

$$\frac{X_1 + \cdots + X_n - n\mu}{\sigma \sqrt{n}} \to \mathcal{N}(0, 1) \text{ as } n \to \infty.$$ 

Also,

$$[A_n - 2 \frac{\sigma}{\sqrt{n}}, A_n + 2 \frac{\sigma}{\sqrt{n}}] \text{ is a 95\% - CI for } \mu.$$
CI for Mean: CLT vs. Chebyshev

Let $X_1, X_2, \ldots$ be i.i.d. with mean $\mu$ and variance $\sigma^2$. Let

$$A_n = \frac{X_1 + \cdots + X_n}{n}.$$ 

The CLT states that

$$\frac{X_1 + \cdots + X_n - n\mu}{\sigma \sqrt{n}} \rightarrow \mathcal{N}(0,1) \text{ as } n \rightarrow \infty.$$ 

Also,

$$[A_n - 2\frac{\sigma}{\sqrt{n}}, A_n + 2\frac{\sigma}{\sqrt{n}}] \text{ is a 95\% CI for } \mu.$$ 

What would Chebyshev’s bound give us?
Cl for Mean: CLT vs. Chebyshev

Let $X_1, X_2, \ldots$ be i.i.d. with mean $\mu$ and variance $\sigma^2$. Let

$$A_n = \frac{X_1 + \cdots + X_n}{n}.$$

The CLT states that

$$\frac{X_1 + \cdots + X_n - n\mu}{\sigma \sqrt{n}} \rightarrow \mathcal{N}(0, 1) \text{ as } n \rightarrow \infty.$$

Also,

$$[A_n - 2 \frac{\sigma}{\sqrt{n}}, A_n + 2 \frac{\sigma}{\sqrt{n}}] \text{ is a 95\% - CI for } \mu.$$ 

What would Chebyshev’s bound give us?

$$[A_n - 4.5 \frac{\sigma}{\sqrt{n}}, A_n + 4.5 \frac{\sigma}{\sqrt{n}}] \text{ is a 95\% - CI for } \mu.$$
CI for Mean: CLT vs. Chebyshev

Let $X_1, X_2, \ldots$ be i.i.d. with mean $\mu$ and variance $\sigma^2$. Let 

$$A_n = \frac{X_1 + \cdots + X_n}{n}.$$ 

The CLT states that 

$$\frac{X_1 + \cdots + X_n - n\mu}{\sigma \sqrt{n}} \rightarrow \mathcal{N}(0, 1) \text{ as } n \rightarrow \infty.$$ 

Also, 

$$[A_n - 2 \frac{\sigma}{\sqrt{n}}, A_n + 2 \frac{\sigma}{\sqrt{n}}] \text{ is a } 95\% \text{- CI for } \mu.$$ 

What would Chebyshev’s bound give us?

$$[A_n - 4.5 \frac{\sigma}{\sqrt{n}}, A_n + 4.5 \frac{\sigma}{\sqrt{n}}] \text{ is a } 95\% \text{- CI for } \mu. (Why?)$$
Cl for Mean: CLT vs. Chebyshev

Let $X_1, X_2, \ldots$ be i.i.d. with mean $\mu$ and variance $\sigma^2$. Let

$$A_n = \frac{X_1 + \cdots + X_n}{n}.$$ 

The CLT states that

$$\frac{X_1 + \cdots + X_n - n\mu}{\sigma \sqrt{n}} \to \mathcal{N}(0, 1) \text{ as } n \to \infty.$$ 

Also,

$$[A_n - 2\frac{\sigma}{\sqrt{n}}, A_n + 2\frac{\sigma}{\sqrt{n}}] \text{ is a 95\% - CI for } \mu.$$ 

What would Chebyshev’s bound give us?

$$[A_n - 4.5\frac{\sigma}{\sqrt{n}}, A_n + 4.5\frac{\sigma}{\sqrt{n}}] \text{ is a 95\% - CI for } \mu. \text{(Why?)}$$ 

Thus, the CLT provides a smaller confidence interval.
Coins and CLT.

Let $X_1, X_2, \ldots$ be i.i.d. $B(p)$. Thus, $X_1 + \cdots + X_n = B(n, p)$.

Here, $\mu = p$ and $\sigma = \sqrt{p(1-p)}$.

CLT states that $X_1 + \cdots + X_n - np \sqrt{p(1-p)} n \to N(0, 1)$. 
Coins and CLT.

Let $X_1, X_2, \ldots$ be i.i.d. $\text{B}(p)$. 
Coins and CLT.

Let $X_1, X_2, \ldots$ be i.i.d. $B(p)$. Thus, $X_1 + \cdots + X_n = B(n, p)$. 

Coins and CLT.

Let $X_1, X_2, \ldots$ be i.i.d. $B(p)$. Thus, $X_1 + \cdots + X_n = B(n, p)$.

Here, $\mu = p$
Coins and CLT.

Let $X_1, X_2, \ldots$ be i.i.d. $B(p)$. Thus, $X_1 + \cdots + X_n = B(n, p)$.

Here, $\mu = p$ and $\sigma = \sqrt{p(1 - p)}$. 
Coins and CLT.

Let $X_1, X_2, \ldots$ be i.i.d. $B(p)$. Thus, $X_1 + \cdots + X_n = B(n, p)$. Here, $\mu = p$ and $\sigma = \sqrt{p(1-p)}$. CLT states that

$$\frac{X_1 + \cdots + X_n - np}{\sqrt{p(1-p)n}} \to \mathcal{N}(0, 1).$$
Coins and CLT.

Let $X_1, X_2, \ldots$ be i.i.d. $B(p)$. Thus, $X_1 + \cdots + X_n = B(n, p)$.

Here, $\mu = p$ and $\sigma = \sqrt{p(1-p)}$. CLT states that

$$\frac{X_1 + \cdots + X_n - np}{\sqrt{p(1-p)n}} \to \mathcal{N}(0, 1).$$
Coins and CLT.

Let $X_1, X_2, \ldots$ be i.i.d. $B(p)$. Thus, $X_1 + \cdots + X_n = B(n, p)$. Here, $\mu = p$ and $\sigma = \sqrt{p(1-p)}$. CLT states that

$$\frac{X_1 + \cdots + X_n - np}{\sqrt{p(1-p)n}} \rightarrow \mathcal{N}(0,1)$$

Thus, $\left[ A_n - 2\sigma \sqrt{n}, A_n + 2\sigma \sqrt{n} \right]$ is a 95\% CI for $\mu$. Since $\sigma \leq 0.5$, $\left[ A_n - 20 \sqrt{n}, A_n + 20 \sqrt{n} \right]$ is a 95\% CI for $\mu$. Therefore, $\left[ A_n - \sqrt{n}, A_n + \sqrt{n} \right]$ is a 95\% CI for $p$. Since $\sigma \leq 0.5$, $\left[ A_n - 20 \sqrt{n}, A_n + 20 \sqrt{n} \right]$ is a 95\% CI for $p$. Thus, $\left[ A_n - 1 \sqrt{n}, A_n + 1 \sqrt{n} \right]$ is a 95\% CI for $p$. 

Coins and CLT.

Let $X_1, X_2, \ldots$ be i.i.d. $B(p)$. Thus, $X_1 + \cdots + X_n = B(n, p)$. Here, $\mu = p$ and $\sigma = \sqrt{p(1-p)}$. CLT states that

$$\frac{X_1 + \cdots + X_n - np}{\sqrt{p(1-p)n}} \rightarrow \mathcal{N}(0, 1)$$

and

$$[A_n - 2 \frac{\sigma}{\sqrt{n}}, A_n + 2 \frac{\sigma}{\sqrt{n}}]$$

is a $95\%$ CI for $\mu$ with $A_n = (X_1 + \cdots + X_n)/n$. 
Coins and CLT.

Let $X_1, X_2, \ldots$ be i.i.d. $B(p)$. Thus, $X_1 + \cdots + X_n = B(n, p)$. Here, $\mu = p$ and $\sigma = \sqrt{p(1 - p)}$. CLT states that

$$\frac{X_1 + \cdots + X_n - np}{\sqrt{p(1 - p)n}} \to \mathcal{N}(0, 1)$$

and

$$[A_n - 2 \frac{\sigma}{\sqrt{n}}, A_n + 2 \frac{\sigma}{\sqrt{n}}]$$

is a 95% – CI for $\mu$

with $A_n = (X_1 + \cdots + X_n)/n$.

Hence,

$$[A_n - 2 \frac{\sigma}{\sqrt{n}}, A_n + 2 \frac{\sigma}{\sqrt{n}}]$$

is a 95% – CI for $p$. 
Coins and CLT.

Let \(X_1, X_2, \ldots\) be i.i.d. \(B(p)\). Thus, \(X_1 + \cdots + X_n = B(n, p)\).

Here, \(\mu = p\) and \(\sigma = \sqrt{p(1-p)}\). CLT states that

\[
\frac{X_1 + \cdots + X_n - np}{\sqrt{np(1-p)n}} \to \mathcal{N}(0, 1)
\]

and

\[
[A_n - 2 \frac{\sigma}{\sqrt{n}}, A_n + 2 \frac{\sigma}{\sqrt{n}}] \text{ is a 95\% CI for } \mu
\]

with \(A_n = (X_1 + \cdots + X_n)/n\).

Hence,

\[
[A_n - 2 \frac{\sigma}{\sqrt{n}}, A_n + 2 \frac{\sigma}{\sqrt{n}}] \text{ is a 95\% CI for } p.
\]

Since \(\sigma \leq 0.5\),
Coins and CLT.

Let $X_1, X_2, \ldots$ be i.i.d. $B(p)$. Thus, $X_1 + \cdots + X_n = B(n, p)$. Here, $\mu = p$ and $\sigma = \sqrt{p(1 - p)}$. CLT states that

$$\frac{X_1 + \cdots + X_n - np}{\sqrt{p(1 - p)n}} \to \mathcal{N}(0, 1)$$

and

$$[A_n - 2 \frac{\sigma}{\sqrt{n}}, A_n + 2 \frac{\sigma}{\sqrt{n}}] \text{ is a 95\% CI for } \mu$$

with $A_n = (X_1 + \cdots + X_n)/n$.

Hence,

$$[A_n - 2 \frac{\sigma}{\sqrt{n}}, A_n + 2 \frac{\sigma}{\sqrt{n}}] \text{ is a 95\% CI for } p.$$

Since $\sigma \leq 0.5$,

$$[A_n - 2 \frac{0.5}{\sqrt{n}}, A_n + 2 \frac{0.5}{\sqrt{n}}] \text{ is a 95\% CI for } p.$$
Coins and CLT.

Let $X_1, X_2, \ldots$ be i.i.d. $B(p)$. Thus, $X_1 + \cdots + X_n = B(n,p)$. Here, $\mu = p$ and $\sigma = \sqrt{p(1-p)}$. CLT states that

$$\frac{X_1 + \cdots + X_n - np}{\sqrt{p(1-p)n}} \to \mathcal{N}(0,1)$$

and

$$[A_n - 2 \frac{\sigma}{\sqrt{n}}, A_n + 2 \frac{\sigma}{\sqrt{n}}]$$

is a 95% - CI for $\mu$ with $A_n = (X_1 + \cdots + X_n)/n$.

Hence,

$$[A_n - 2 \frac{\sigma}{\sqrt{n}}, A_n + 2 \frac{\sigma}{\sqrt{n}}]$$

is a 95% - CI for $p$.

Since $\sigma \leq 0.5$,

$$[A_n - 2 \frac{0.5}{\sqrt{n}}, A_n + 2 \frac{0.5}{\sqrt{n}}]$$

is a 95% - CI for $p$.

Thus,

$$[A_n - \frac{1}{\sqrt{n}}, A_n + \frac{1}{\sqrt{n}}]$$

is a 95% - CI for $p$. 

Comparing Chebyshev and CLT: Polling

We ask $n$ randomly sampled voters whether they support Bob. $X_i = 1$ if the $i^{th}$ voter says "yes" and $X_i = 0$ otherwise. The $X_i$ are iid.

We want to be sure with prob $\geq 0.95$ that $|M_{100} - p| \leq 0.1$. How many people should we ask?

Again, use the bound that $\text{var}(X_i) \leq \frac{1}{4}$.

By Chebyshev:

$$\frac{25}{n} \leq 0.05 \Rightarrow n \geq 500$$

By CLT:

$$2(1 - \phi(2 \times 0.1 \times \sqrt{n})) \leq 0.05$$

$$\phi(2 \times 0.1 \times \sqrt{n}) \geq 0.975$$

Since $\phi(1.96) = 0.975$:

$$n \geq 96.04$$

CLT much better than Chebyshev.
Inequalities and Confidence Intervals

1. Inequalities: Markov and Chebyshev Tail Bounds
2. Weak Law of Large Numbers
3. Confidence Intervals: Chebyshev Bounds vs. CLT Approx.
4. CLT: $X_n$ i.i.d. $\implies \frac{A_n - \mu}{\sigma/\sqrt{n}} \to \mathcal{N}(0, 1)$
Summary

Inequalities and Confidence Interals

1. Inequalities: Markov and Chebyshev Tail Bounds
2. Weak Law of Large Numbers
3. Confidence Intervals: Chebyshev Bounds vs. CLT Approx.
4. CLT: \( X_n \) i.i.d. \( \implies \frac{A_n - \mu}{\sigma / \sqrt{n}} \to \mathcal{N}(0, 1) \)
5. CI: \([A_n - 2 \frac{\sigma}{\sqrt{n}}, A_n + 2 \frac{\sigma}{\sqrt{n}}]\) = 95\%-CI for \( \mu \).
Summary

Inequalities and Confidence Intervals

1. Inequalities: Markov and Chebyshev Tail Bounds
2. Weak Law of Large Numbers
3. Confidence Intervals: Chebyshev Bounds vs. CLT Approx.
4. CLT: $X_n \text{ i.i.d.} \implies \frac{A_n - \mu}{\sigma / \sqrt{n}} \to \mathcal{N}(0,1)$
5. CI: $[A_n - 2 \frac{\sigma}{\sqrt{n}}, A_n + 2 \frac{\sigma}{\sqrt{n}}] = 95\%-\text{CI for } \mu$. 