Concentration Inequalities

Shusen Wang

Bernstein Inequality

Bernstein Inequality

- Let $Z_1, \cdots, Z_n$ be independent zero-mean random variables.
- They are bounded: $|Z_i| \leq b$, for all i.
- Their variances are bounded: $\mathbb{E}\left[Z_i^2\right] \leq v$, for all i.

Bernstein Inequality

- Let $Z_1, \cdots, Z_n$ be independent zero-mean random variables.
- They are bounded: $|Z_i| \le b$, for all i.
- Their variances are bounded: $\mathbb{E}\left[Z_i^2\right] \le v$, for all i.
- Bernstein inequality:

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| \ \ge \ \epsilon\right) \quad \le \quad \exp\left(-\frac{\epsilon^2 n/2}{v + \epsilon b/3}\right).$$

Bernstein Inequality

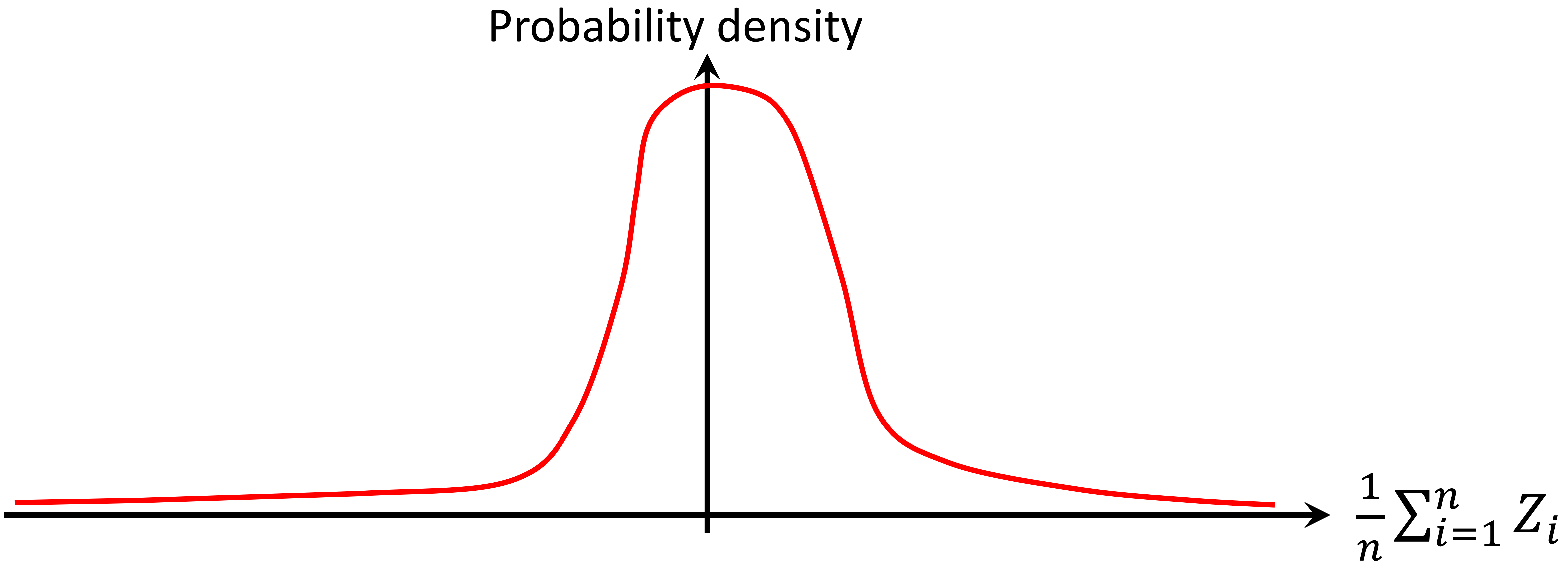
- Let $Z_1, \cdots, Z_n$ be independent zero-mean random variables.
- They are bounded: $|Z_i| \leq b$, for all i.
- Their variances are bounded: $\mathbb{E}\left[Z_i^2\right] \leq v$, for all i.
- Bernstein inequality:

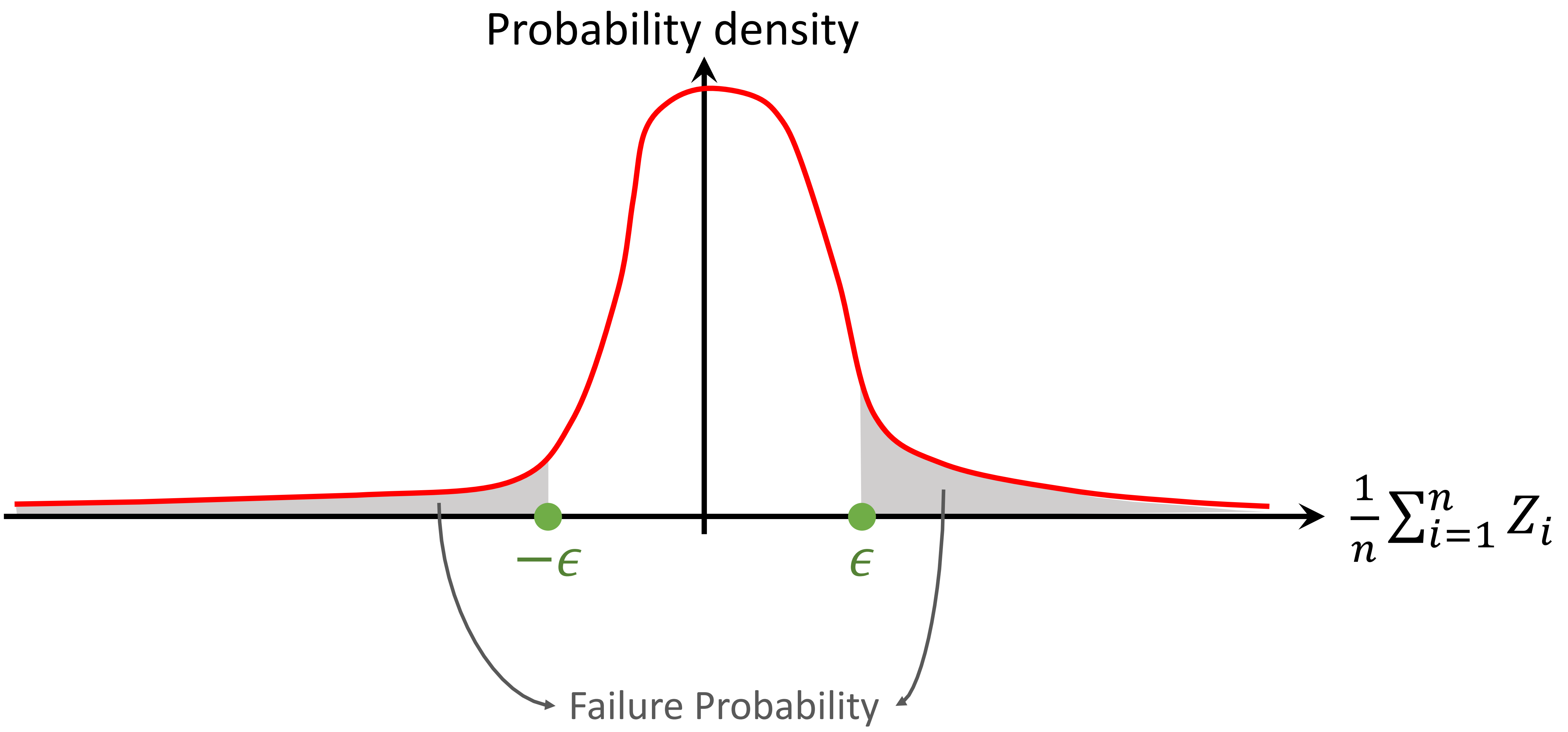
$$\mathbb{P}\left(\left|\underbrace{\frac{1}{n}\sum_{i=1}^{n} Z_i}_{\text{Mean}}\right| \geq \epsilon\right) \leq \exp\left(-\frac{\epsilon^2 n/2}{v + \epsilon b/3}\right).$$

Bernstein Inequality

- Let $Z_1, \cdots, Z_n$ be independent zero-mean random variables.
- They are bounded: $|Z_i| \leq b$, for all i.
- Their variances are bounded: $\mathbb{E}\left[Z_i^2\right] \leq v$, for all i.
- Bernstein inequality:

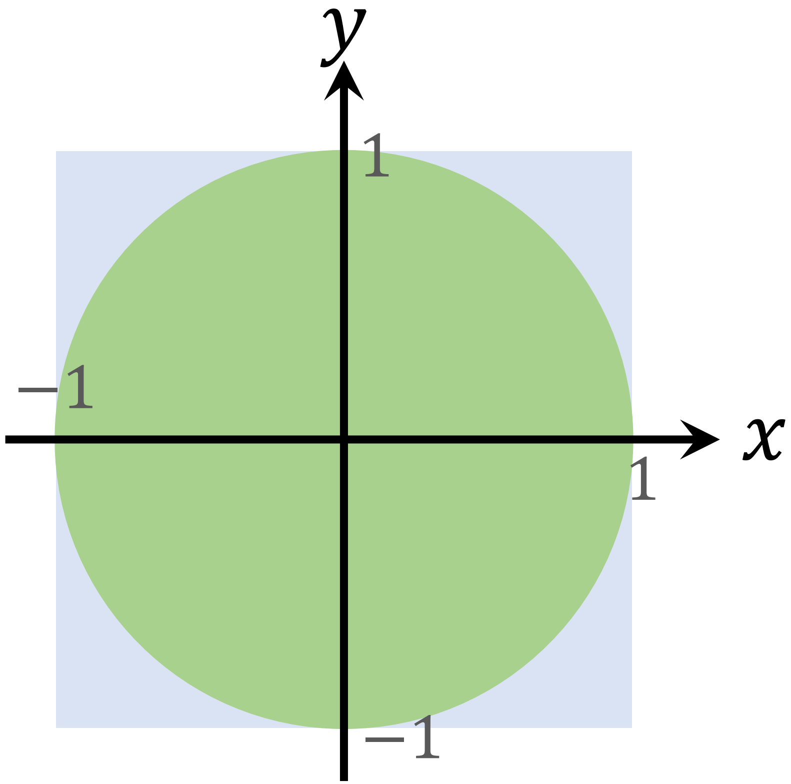
$$\underbrace{\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| \geq \epsilon\right)}_{\text{Failure Probability}} \leq \underbrace{\exp\left(-\frac{\epsilon^2 n/2}{v+\epsilon b/3}\right)}_{\text{Decays exponentially with } n}.$$


Probability density
$\frac{1}{n}\sum_{i=1}^{n} Z_i$


Probability density
$\frac{1}{n}\sum_{i=1}^{n} Z_i$
$-\epsilon$
ϵ
Failure Probability

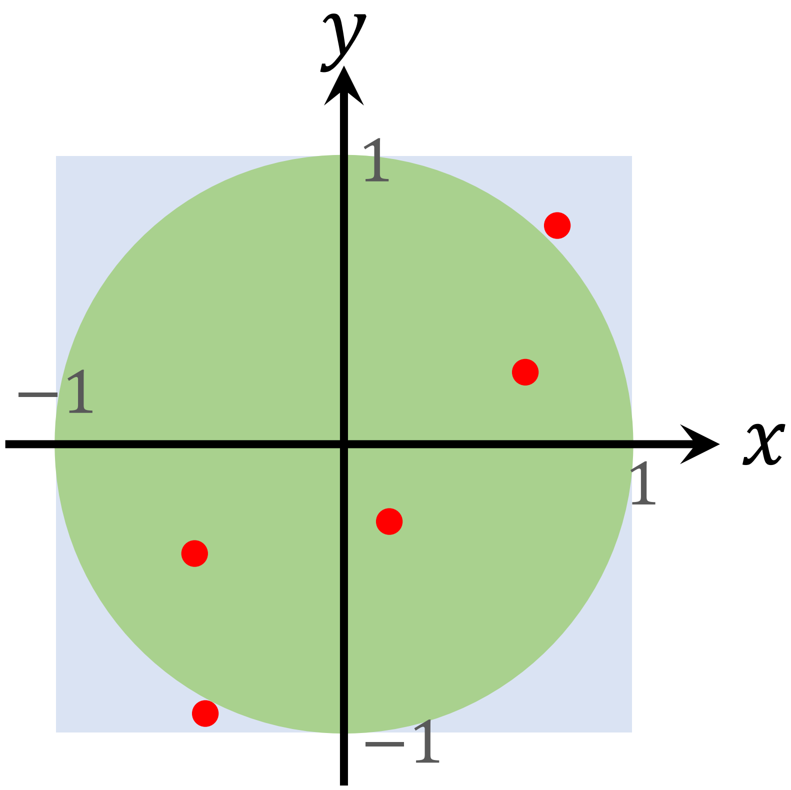
Approximating Pi Using Monte Carlo

Monte Carlo Approximation to Pi



- Let $f(x, y) = \begin{cases} 4, \text{if } x^2 + y^2 \leq 1; \\ 0, \quad \text{otherwise.} \end{cases}$

Monte Carlo Approximation to Pi



- Let $f(x, y) = \begin{cases} 4, \text{if } x^2 + y^2 \leq 1; \\ 0, \quad \text{otherwise.} \end{cases}$
- Draw n samples from the square uniformly at random, denote

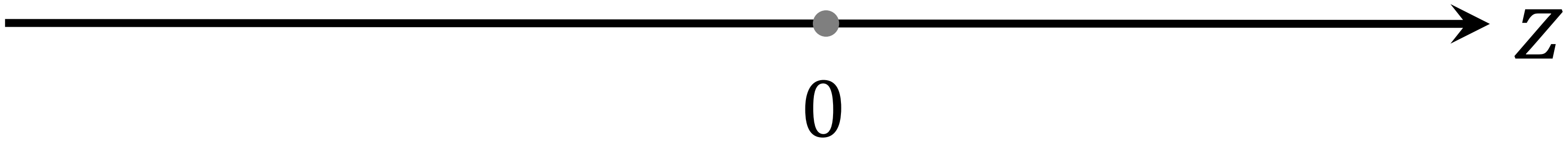
$$(x_1, y_1), \cdots, (x_n, y_n).$$

- Calculate $Q_n = \frac{1}{n} \sum_{i=1}^{n} f(x_i, y_i)$.
- $\mathbb{E}[Q_n] = \pi$.

Analyzing Convergence Rate

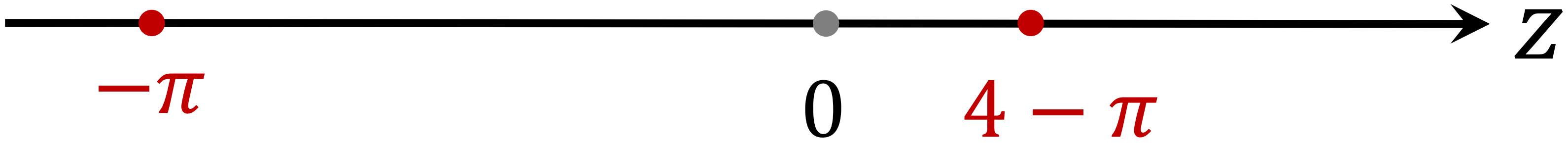
Applying Bernstein Inequality

- Define random variable $Z = f(X, Y) - \pi$.
- Z has zero mean: $\mathbb{E}[Z] = 0$. (Because $\mathbb{E}[f(X, Y)] = \pi$.)



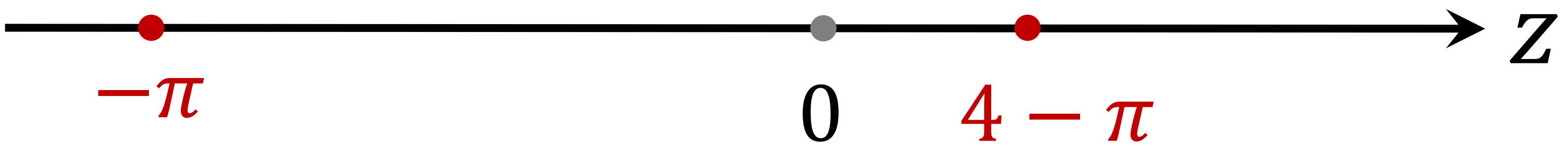
Applying Bernstein Inequality

- Define random variable $Z = f(X, Y) - \pi$.
- Z has zero mean: $\mathbb{E}[Z] = 0$. (Because $\mathbb{E}[f(X,Y)] = \pi$.)
- $Z \in \{-\pi, 4 - \pi\}$. (Because $f(X, Y)$ is either 0 or 4.)



Applying Bernstein Inequality

- Define random variable $Z = f(X,Y) - \pi$.
- Z has zero mean: $\mathbb{E}[Z] = 0$. (Because $\mathbb{E}[f(X,Y)] = \pi$.)
- $Z \in \{-\pi, 4-\pi\}$. (Because $f(X,Y)$ is either 0 or 4.)
- Z is bounded: $|Z| \le \pi$.
- The variance is bounded: $\mathbb{E}[Z^2] \le \pi^2$.



Applying Bernstein Inequality

- $Z_i = f(X_i, Y_i) - \pi$ is a zero-mean random variables.
- Bernstein inequality:

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| \geq \epsilon\right) \quad \leq \quad \exp\left(-\frac{\epsilon^2 n/2}{v + \epsilon b/3}\right).$$

- We found $b = \pi$ and $v = \pi^2$.

Applying Bernstein Inequality

- $Z_i = f(X_i, Y_i) - \pi$ is a zero-mean random variables.
- Bernstein inequality:

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| \geq \epsilon\right) \quad \leq \quad \exp\left(-\frac{\epsilon^2 n/2}{v + \epsilon b/3}\right).$$

- We found $b = \pi$ and $v = \pi^2$.

Lemma 1: $\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| \geq \epsilon\right) \leq \exp\left(-\frac{\epsilon^2 n/2}{\pi^2 + \epsilon\pi/3}\right).$

Establishing Convergence Rate

Lemma 1: $\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| \geq \epsilon\right) \leq \boxed{\exp\left(-\frac{\epsilon^2 n/2}{\pi^2 + \epsilon\pi/3}\right)}.$

$\triangleq \delta$

Establishing Convergence Rate

Lemma 1: $\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| \geq \epsilon\right) \leq \boxed{\exp\left(-\frac{\epsilon^2 n/2}{\pi^2+\epsilon\pi/3}\right)}.$

$\triangleq \delta$

- $\delta \triangleq \exp\left(-\frac{\epsilon^2 n/2}{\pi^2+\epsilon\pi/3}\right)$

Establishing Convergence Rate

Lemma 1: $\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| \geq \boxed{\epsilon}\right) \leq \boxed{\exp\left(-\frac{\epsilon^2 n/2}{\pi^2 + \epsilon\pi/3}\right)}.$

$= \frac{4.7}{\sqrt{n}} \cdot \ln^{\frac{1}{2}}\left(\frac{1}{\delta}\right)$ $\triangleq \delta$

- $\delta \triangleq \exp\left(-\frac{\epsilon^2 n/2}{\pi^2 + \epsilon\pi/3}\right)$ ➔ $\epsilon = \frac{4.7}{\sqrt{n}} \cdot \ln^{\frac{1}{2}}\left(\frac{1}{\delta}\right).$

Establishing Convergence Rate

Lemma 1: $\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| \geq \boxed{\epsilon}\right) \leq \boxed{\exp\left(-\frac{\epsilon^2 n/2}{\pi^2 + \epsilon\pi/3}\right)}.$

$$= \frac{4.7}{\sqrt{n}} \cdot \ln^{\frac{1}{2}}\left(\frac{1}{\delta}\right) \qquad \triangleq \delta$$

Lemma 2: $\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| \geq \frac{4.7}{\sqrt{n}} \cdot \ln^{\frac{1}{2}}\left(\frac{1}{\delta}\right)\right) \leq \delta.$

Establishing Convergence Rate

Lemma 2: $\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| \geq \frac{4.7}{\sqrt{n}} \cdot \ln^{\frac{1}{2}}\left(\frac{1}{\delta}\right)\right) \leq \delta.$

Lemma 2': $\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| \leq \frac{4.7}{\sqrt{n}} \cdot \ln^{\frac{1}{2}}\left(\frac{1}{\delta}\right)$ holds w.p. at least $1-\delta$.

Establishing Convergence Rate

Lemma 2: $\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| \geq \frac{4.7}{\sqrt{n}} \cdot \ln^{\frac{1}{2}}\left(\frac{1}{\delta}\right)\right) \leq \delta.$

$$\Downarrow$$

Lemma 2': $\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| \leq \frac{4.7}{\sqrt{n}} \cdot \ln^{\frac{1}{2}}\left(\frac{1}{\delta}\right)$ holds w.p. at least $1-\delta$.

- By definition, $Z_i = f(X_i, Y_i) - \pi$ and $Q_n = \frac{1}{n}\sum_{i=1}^{n} f(X_i, Y_i)$.

Establishing Convergence Rate

Lemma 2: $\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| \geq \frac{4.7}{\sqrt{n}} \cdot \ln^{\frac{1}{2}}\left(\frac{1}{\delta}\right)\right) \leq \delta.$

$\Downarrow$

Lemma 2': $\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| \leq \frac{4.7}{\sqrt{n}} \cdot \ln^{\frac{1}{2}}\left(\frac{1}{\delta}\right)$ holds w.p. at least $1-\delta$.

- By definition, $Z_i = f(X_i, Y_i) - \pi$ and $Q_n = \frac{1}{n}\sum_{i=1}^{n} f(X_i, Y_i)$.
- $\Rightarrow$ $\frac{1}{n}\sum_{i=1}^{n} Z_i = \frac{1}{n}\sum_{i=1}^{n} f(X_i, Y_i) - \pi$

Establishing Convergence Rate

Lemma 2: $\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| \geq \frac{4.7}{\sqrt{n}} \cdot \ln^{\frac{1}{2}}\left(\frac{1}{\delta}\right)\right) \leq \delta.$

⬇

Lemma 2': $\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| \leq \frac{4.7}{\sqrt{n}} \cdot \ln^{\frac{1}{2}}\left(\frac{1}{\delta}\right)$ holds w.p. at least $1-\delta$.

- By definition, $Z_i = f(X_i, Y_i) - \pi$ and $Q_n = \frac{1}{n}\sum_{i=1}^{n} f(X_i, Y_i)$.
- ➔ $\frac{1}{n}\sum_{i=1}^{n} Z_i = \boxed{\frac{1}{n}\sum_{i=1}^{n} f(X_i, Y_i)} - \pi$

$\triangleq Q_n$

Establishing Convergence Rate

Lemma 2: $\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| \geq \frac{4.7}{\sqrt{n}} \cdot \ln^{\frac{1}{2}}\left(\frac{1}{\delta}\right)\right) \leq \delta.$

Lemma 2': $\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| \leq \frac{4.7}{\sqrt{n}} \cdot \ln^{\frac{1}{2}}\left(\frac{1}{\delta}\right)$ holds w.p. at least $1-\delta$.

- By definition, $Z_i = f(X_i, Y_i) - \pi$ and $Q_n = \frac{1}{n}\sum_{i=1}^{n} f(X_i, Y_i)$.
- ➔ $\frac{1}{n}\sum_{i=1}^{n} Z_i = \frac{1}{n}\sum_{i=1}^{n} f(X_i, Y_i) - \pi = \underline{Q_n - \pi}.$

Establishing Convergence Rate

Lemma 2: $\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| \geq \frac{4.7}{\sqrt{n}} \cdot \ln^{\frac{1}{2}}\left(\frac{1}{\delta}\right)\right) \leq \delta.$

$\Downarrow$

Lemma 2’: $\boxed{\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right|} \leq \frac{4.7}{\sqrt{n}} \cdot \ln^{\frac{1}{2}}\left(\frac{1}{\delta}\right)$ holds w.p. at least $1-\delta$.

$\triangleq |Q_n - \pi|$

Establishing Convergence Rate

Lemma 2: $\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| \geq \frac{4.7}{\sqrt{n}} \cdot \ln^{\frac{1}{2}}\left(\frac{1}{\delta}\right)\right) \leq \delta.$

$\Downarrow$

Lemma 2': $\boxed{\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right|} \leq \frac{4.7}{\sqrt{n}} \cdot \ln^{\frac{1}{2}}\left(\frac{1}{\delta}\right)$ holds w.p. at least $1-\delta$.

$\triangleq |Q_n - \pi|$

$\Downarrow$

Theorem: $\underline{|Q_n - \pi|} \leq \underline{\frac{4.7}{\sqrt{n}} \cdot \ln^{\frac{1}{2}}\left(\frac{1}{\delta}\right)}$ holds w.p. at least $1-\delta$.

Thank you!