Topics in Statistics: Markov chain Monte Carlo
Problem Set 0 - Solutions

Exercise 1

Part 1: Prove Markov's Inequality

Proof. Let Z be a positive random variable with $\mathbb{E}[Z] < \infty$ and let $a > 0$. Following the hint, we can write:

$$\mathbb{P}(Z > a) = \mathbb{E}[\mathbf{1}\{Z > a\}]$$

Now observe that for all values of Z:

- When $Z > a$: $\mathbf{1}\{Z > a\} = 1 \leq \frac{Z}{a}$
- When $Z \leq a$: $\mathbf{1}\{Z > a\} = 0 \leq \frac{Z}{a}$

Therefore, $\mathbf{1}\{Z > a\} \leq \frac{Z}{a}$ for all values of Z.
Taking expectations on both sides:

$$\mathbb{E}[\mathbf{1}\{Z > a\}] \leq \mathbb{E}\left[\frac{Z}{a}\right] = \frac{\mathbb{E}[Z]}{a}$$

Thus:

$$\boxed{\mathbb{P}(Z > a) \leq \frac{\mathbb{E}[Z]}{a}}$$

□

Part 2: Prove Chebyshev's Inequality

Proof. Let X be a random variable with $\mathbb{E}[X^2] < \infty$, $\mu := \mathbb{E}[X]$, and $\sigma^2 := \mathrm{Var}(X)$. We need to show that for any $t > 0$:

$$\mathbb{P}[|X - \mu| \geq t] \leq \frac{\sigma^2}{t^2}$$

Define $Z = (X - \mu)^2$. Note that Z is a positive random variable with:

$$\mathbb{E}[Z] = \mathbb{E}[(X - \mu)^2] = \mathrm{Var}(X) = \sigma^2$$

Applying Markov's inequality to Z with $a = t^2$:

$$\mathbb{P}[(X-\mu)^2 > t^2] \leq \frac{\mathbb{E}[(X-\mu)^2]}{t^2} = \frac{\sigma^2}{t^2}$$

Since $(X-\mu)^2 > t^2$ is equivalent to $|X-\mu| > t$, we have:

$$\boxed{\mathbb{P}[|X-\mu| \geq t] \leq \frac{\sigma^2}{t^2}}$$

□

Part 3: Prove the Weak Law of Large Numbers

Proof. Let $X_1, X_2, \ldots$ be i.i.d. random variables with $\mu := \mathbb{E}[X_1]$ and $\sigma^2 := \mathrm{Var}(X_1)$. Define:

$$S_n := \frac{1}{n}\sum_{i=1}^{n} X_i, \quad n \geq 1$$

First, we compute the expectation and variance of S_n:

$$\mathbb{E}[S_n] = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[X_i] = \frac{1}{n} \cdot n \cdot \mu = \mu$$

Using independence of the X_i:

$$\mathrm{Var}(S_n) = \mathrm{Var}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}(X_i) = \frac{1}{n^2} \cdot n \cdot \sigma^2 = \frac{\sigma^2}{n}$$

Applying Chebyshev's inequality to S_n with parameter $\epsilon > 0$:

$$\mathbb{P}[|S_n - \mu| \geq \epsilon] \leq \frac{\mathrm{Var}(S_n)}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2}$$

As $n \to \infty$:

$$\frac{\sigma^2}{n\epsilon^2} \to 0$$

Therefore:

$$\boxed{\mathbb{P}[|S_n - \mu| \geq \epsilon] \to 0 \text{ as } n \to \infty}$$

This proves that the sample mean converges in probability to the expectation. □

Exercise 2

Part 1: Compute $\mathbb{E}[g(X)]$ and Var$[g(X)]$

Proof. The function $g(x) = \mathbf{1}\{x \in A\}$ is the indicator function of the quarter disk A.

Since X is uniformly distributed on the unit square U, the probability that X falls in A is:

$$\mathbb{E}[g(X)] = \mathbb{P}(X \in A) = \frac{\text{Area}(A \cap U)}{\text{Area}(U)} = \frac{\pi/4}{1} = \frac{\pi}{4}$$

For the variance, note that since $g(X) \in \{0, 1\}$, we have $g(X)^2 = g(X)$. Therefore:

$$\mathbb{E}[g(X)^2] = \mathbb{E}[g(X)] = \frac{\pi}{4}$$

The variance is:

$$\text{Var}[g(X)] = \mathbb{E}[g(X)^2] - (\mathbb{E}[g(X)])^2 = \frac{\pi}{4} - \left(\frac{\pi}{4}\right)^2 = \frac{\pi}{4}\left(1 - \frac{\pi}{4}\right) = \frac{\pi(4-\pi)}{16}$$

Thus:

$$\boxed{\mathbb{E}[g(X)] = \frac{\pi}{4}, \quad \text{Var}[g(X)] = \frac{\pi(4-\pi)}{16}}$$

□

Part 2: Construct a Consistent Estimator and Confidence Interval

Proof. Define the estimator:

$$\hat{\pi}_n = 4 \cdot \frac{1}{n}\sum_{i=1}^{n} g(X_i)$$

By the Law of Large Numbers:

$$\frac{1}{n}\sum_{i=1}^{n} g(X_i) \xrightarrow{P} \mathbb{E}[g(X)] = \frac{\pi}{4}$$

Therefore:

$$\hat{\pi}_n \xrightarrow{P} \pi$$

This shows $\hat{\pi}_n$ is a consistent estimator of π.

For the confidence interval, compute:

$$\mathbb{E}[\hat{\pi}_n] = 4 \cdot \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} g(X_i)\right] = 4 \cdot \frac{\pi}{4} = \pi$$

$$\text{Var}(\hat{\pi}_n) = 16 \cdot \text{Var}\left[\frac{1}{n}\sum_{i=1}^{n} g(X_i)\right] = 16 \cdot \frac{1}{n^2} \cdot n \cdot \text{Var}[g(X)] = \frac{16 \cdot \pi(4-\pi)}{16n} = \frac{\pi(4-\pi)}{n}$$

By Chebyshev's inequality:

$$\mathbb{P}(|\hat{\pi}_n - \pi| \geq t) \leq \frac{\text{Var}(\hat{\pi}_n)}{t^2} = \frac{\pi(4-\pi)}{nt^2}$$

For a $(1-\alpha)$ confidence interval, we want $\mathbb{P}(|\hat{\pi}_n - \pi| \leq B_n) \geq 1 - \alpha$. Setting $\frac{\pi(4-\pi)}{nB_n^2} = \alpha$ and solving for B_n:

$$B_n = \sqrt{\frac{\pi(4-\pi)}{n\alpha}}$$

Since π is unknown, we use the upper bound $\pi(4-\pi) \leq 1$ (maximum occurs at $\pi = 2$):

$$B_n = \sqrt{\frac{1}{n\alpha}}$$

The $(1-\alpha)$ confidence interval is:

$$\boxed{\left[\hat{\pi}_n - \sqrt{\frac{1}{n\alpha}}, \hat{\pi}_n + \sqrt{\frac{1}{n\alpha}}\right]}$$

with $A_n = B_n = \sqrt{\frac{1}{n\alpha}}$. □

Part 3: Tighter Confidence Interval Using Higher Moments

Proof. To obtain a tighter confidence interval, we can apply Markov's inequality to $|\hat{\pi}_n - \pi|^k$ for $k > 2$.

For any $t > 0$ and $k > 0$:

$$\mathbb{P}(|\hat{\pi}_n - \pi| \geq t) = \mathbb{P}(|\hat{\pi}_n - \pi|^k \geq t^k) \leq \frac{\mathbb{E}[|\hat{\pi}_n - \pi|^k]}{t^k}$$

Let $Z_i = g(X_i) - \frac{\pi}{4}$, which are i.i.d. with zero mean. Then:

$$\hat{\pi}_n - \pi = \frac{4}{n}\sum_{i=1}^{n} Z_i$$

Using the hint that $\mathbb{E}\left[\left|\sum_{i=1}^n Z_i\right|^k\right] \leq Cn^{k/2}$ for some constant $C > 0$:

$$\mathbb{E}[|\hat{\pi}_n - \pi|^k] = \left(\frac{4}{n}\right)^k \mathbb{E}\left[\left|\sum_{i=1}^{n} Z_i\right|^k\right] \leq \left(\frac{4}{n}\right)^k \cdot Cn^{k/2} = C \cdot 4^k \cdot n^{-k/2}$$

Therefore:

$$\mathbb{P}(|\hat{\pi}_n - \pi| \geq t) \leq \frac{C \cdot 4^k \cdot n^{-k/2}}{t^k}$$

For a $(1-\alpha)$ confidence interval, set this equal to α:

$$t = \left(\frac{C \cdot 4^k}{\alpha \cdot n^{k/2}}\right)^{1/k}$$

The confidence interval width is:

$$2t = 2\left(\frac{C \cdot 4^k}{\alpha \cdot n^{k/2}}\right)^{1/k} = O(n^{-1/2})$$

As k increases, the constant improves but the asymptotic rate remains $O(n^{-1/2})$. This provides a tighter confidence interval than the one obtained using only second moments, especially for large n.

Alternative approach: For even tighter bounds, one could use the Central Limit Theorem (for large n) or Hoeffding's inequality (since $g(X_i)$ is bounded), which would give exponentially decaying tail probabilities rather than polynomial decay. $\square$