CS215 Assignment1 Problem 3

Anish Kulkarni and Parth Pujari

August 2022

1 Problem

Consider a random variable X. Consider a data set that comprises N independent draws (e.g., modeled by $X_1, \cdots, X_N$) from the distribution of X.
Use the law of large numbers to show that the random variable $\hat{M} = \frac{(X_1+\cdots+X_N)}{N}$ converges to the true mean $M = E(X)$ as $N \to \infty$
Prove that the expected value of the random variable $\hat{V} = \sum_{i=1}^{N} \frac{(X_i - \hat{M})^2}{N}$ tends to the true variance $V = Var(X)$ as $N \to \infty$

Proof

Part 1

Let x be any real number other than $M = E[X]$. Consider $P(x \leq \hat{M} \leq x + dx)$.
Choose an ϵ such that $|x - M| > 2\epsilon$
Now by Law of Large Numbers $P(|\hat{M} - M| \geq \epsilon) \to 0$ as $n \to \infty$.
But due to choice of ϵ

$$P(x \leq \hat{M} \leq x + dx) \leq P(|\hat{M} - M| \geq \epsilon)$$

And hence $P(x \leq \hat{M} \leq x + dx) \to 0$ as $n \to \infty$
As this holds for all $x \neq M$ we have that $\hat{M}$ converges to the true mean M
Hence proved.

Part 2

$$\hat{V} = \sum_{i=1}^{N} \frac{(X_i - \hat{M})^2}{N}$$

$$= \sum_{i=1}^{N} \frac{X_i^2 - 2X_i\hat{M} + \hat{M}^2}{N}$$

$$= \frac{\sum_{i=1}^{N} X_i^2}{N} - 2\hat{M}^2 + \hat{M}^2$$

$$= \frac{\sum_{i=1}^{N} X_i^2}{N} - \hat{M}^2$$

Now by previous result $\hat{M} \to E[X]$ and hence $\hat{M}^2 \to E[X]^2$ as $n \to \infty$.
Applying previous result on X_i^2 gives $\frac{\sum_{i=1}^{N} X_i^2}{N} \to E[X^2]$ as $n \to \infty$.
Hence

$$\hat{V} \to E[X^2] - E[X]^2 = Var(X) \text{ as } n \to \infty$$

Hence proved.

2 Random Walks

walksGraph generates 10000 random walks each with 1000 steps of step length 0.001. The various paths taken and the histogram of the final locations of the walks is shown in the figures below.

3 Empirically Computed values

Mean	Variance
1.386×10^{-4}	9.905×10^{-4}

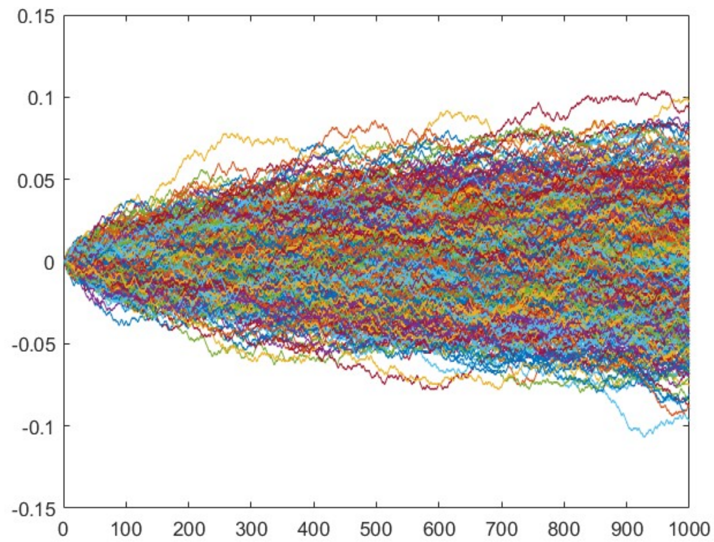

Figure 1: Random Walks

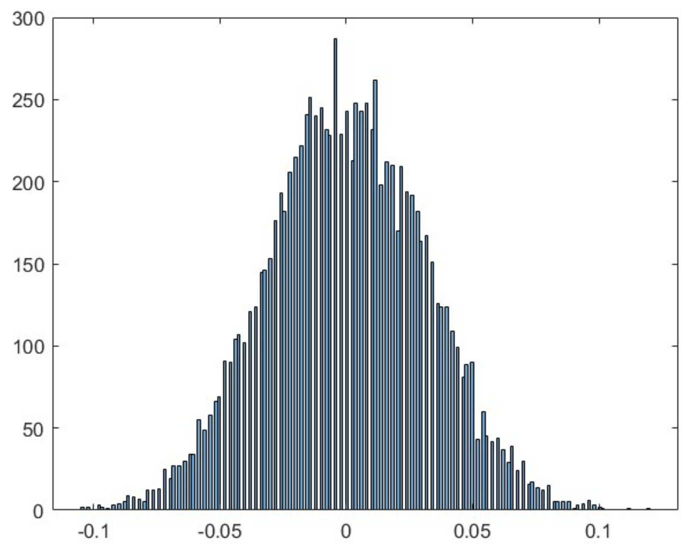

Figure 2: Histogram

4 Theoretical values

Let the random walk experiment have N steps of size e.
Let the random variable X denote the final position of the walker.
We shall calculate $P(X = ke)$, where $-N \leq k \leq N$.
Assume he went forward u times and backward v times, then

$$u + v = N$$

$$ue - ve = ke$$

Solving we get $u = \frac{N+k}{2}$. (In particular k and N must have same parity)
Now we can choose which steps to go forward in $\binom{N}{u}$ ways. Hence

$$P(X = ke) = \binom{N}{\frac{N+k}{2}}$$

4.1 Mean

We have $P(X = k) = P(X = -k)$, hence

$$\text{True Mean } (\hat{M}) = 0$$

4.2 Variance

$$\text{Var } (X) = \sum_{k=-N, k \equiv N mod 2}^{k=N} P(X = k)(ke - 0)^2$$

$$= \sum_{i=0}^{N} P(X = -N + 2i)(N - 2i)^2 e^2$$

$$= e^2 \sum_{i=0}^{N} \binom{N}{i} (N - 2i)^2$$

This can be evaluated using standard formulae and we get

$$\text{Var } (X) = Ne^2$$

Under the given constraints of $N = 1000$ and $e = 0.001$, we get variance = 10^{-3}

5 Error between Theoretical and Empirical

5.1 Mean

$$\text{Err(Mean)} = |\text{ Empirical Mean - True Mean }|$$

$$= |1.386 \times 10^{-4} - 0| = 1.386 \times 10^{-4}$$

5.2 Variance

$$\text{Err(Variance)} = |\text{ Empirical Variance - True Variance }|$$

$$= |9.905 \times 10^{-4} - 0.001| = 9.5 \times 10^{-6}$$