STAT 753: Stochastic Models and Simulations
Jaleesa Houle
University of Nevada, Reno - Spring 2024

Homework 6

Theory

1. Take a random walk $S_0, S_1, \ldots$ starting from zero with independent identically distributed Laplace increments with mean μ and standard deviation σ. Let $M_n := e^{S_n}$ be the corresponding geometric random walk. Find the mean, median, and standard deviation of M_n.

Solution

We will first start by computing $E[S_n]$ and $var(S_n)$. For a Laplace random variable with $E[x] = \mu$ and $var(X) = \sigma^2$, the arithmetic random walk with Z_k IID can be found as

$$\begin{aligned} E[S_n] &= E[S_0 + Z_1 + ...Z_n] \\ &= E[S_0] + E[Z_1] + ...E[Z_n] \\ &= S_0 + \mu + ...\mu. \qquad &= S_0 + n\mu. \end{aligned} \tag{1}$$

$$\begin{aligned} var[S_n] &= var[S_0 + Z_1 + ...Z_n] \\ &= var[S_0] + var[Z_1] + ...var[Z_n] \\ &= 0 + \sigma^2 + ...\sigma^2 \\ &= 0 + n\sigma^2. \end{aligned} \tag{2}$$

So, $S_n \sim L(n\mu, 2\sigma^2)$. For a distribution with $\mu = 0$ and $\sigma^2 = 1$, this reduces to $S_n \sim L(0, n)$. The median for a Laplace random variable with mean μ and variance σ^2 is μ. Therefore, for $S_n = S_0 + Z_1 + ...Z_n$, the median should equal the mean for distribution S_n, and e^μ for distribution M_n. So, in the case of $\mu = 0$, the median of S_n should be 0, and the median of $M_n = e^0 = 1$.

Now, The expected value of $M_n = E[e^{tS_n}]$ can be described by the MGF

$$M_X^i(t) = \frac{e^{\mu t}}{1 - \frac{\sigma^2}{2}t^2}, \text{ for } |t| < \frac{\sqrt{2}}{\sigma}. \tag{3}$$

$$\begin{aligned} E[M_n] &= E[e^{tS_n}]\Big|_{t=1} \\ &= \frac{e^\mu}{1 - \frac{\sigma^2}{2}} \end{aligned} \tag{4}$$

By substituting the values for $E[S_n]$ and $var(S_n)$ found above,

$$E[M_n] = \frac{e^{S_0+n\mu}}{1-\frac{n\sigma^2}{2}}. \tag{5}$$

Similarly,

$$\begin{aligned} E[M_n^2] &= E[e^{tS_n}]\Big|_{t=2} \\ &= \frac{e^{2\mu}}{1-2\sigma^2} \\ &= \frac{e^{2n\mu+2S_0}}{1-2n\sigma^2}. \end{aligned} \tag{6}$$

$$\begin{aligned} var(M_n) &= E[M_n^2] - E[M_n]^2 \\ &= \frac{e^{2n\mu+2S_0}}{1-2n\sigma^2} - \left(\frac{e^{S_0+n\mu}}{1-\frac{n\sigma^2}{2}}\right)^2 \\ &= e^{2S_0+2n\mu}\left(\frac{1}{1-2n\sigma^2} - \frac{1}{(1-\frac{n\sigma^2}{2})^2}\right) \end{aligned} \tag{7}$$

It should be noted that for this MGF, the constraint of $|t| < \sqrt{2}/\sigma$ means that the mean and variance of a random walk M_n will not converge for $|t| > \sqrt{2}/\sigma$. For the coding part of this assignment, I will be using the parameters $\mu = 0$ and $\sigma^2 = 1$, which violates the constraint $|t| < \sqrt{2}/\sigma$ for $E[M_n^2]$. Thus the theoretical mean and variance for M_n will not converge, but the median should still be equal to 1.

2. For a simple random walk $(S_0, S_1, ...)$ with probabilities 0.6, 0.4 going up or down by 1, starting from $S_0 = -1$, find the probability that:

a) at time 12 the position is $S_{12} = 5$.

Solution

The probability of a simple random walk being having a position of 5 at step 12, given a starting position of -1 can be expressed as $P(S_{12} = 5|S_1 = -1)$. The solutions to this transition probability is $\binom{n}{a} p^a q^b$, where n is the number of steps, p is the probability of going up by 1, q is the probability of going down by 1, a is the number of steps up, and b is the number of steps down. This expression must satisfy $a + b = n$ and $a - b = z$, where z is the altitude change. In this case, the altitude change from -1 to 5 is $z = 6$ and the number of steps is $n = 12$. Solving this system of equations, we find that $a = 9$ and $b = 3$. Thus,

$$\begin{aligned} P(S_{12} = 5|S_1 = -1) &= \binom{12}{9} 0.6^9 0.4^3 \\ &= 0.1419. \end{aligned} \tag{8}$$

b) $P(S_{12} = 5|S_4 = 3)$.

Solution

The probability $P(S_{12} = 5|S_4 = 3)$ can be found in the same way as above. In this case, we can consider the starting position to be 3, and the number of steps to be $12 - 4 = 8$. The change in altitude is then $5 - 3 = 2$. Therefore, we can again solve the system of equations $a + b = 8$ and $a - b = 2$. Doing so results in $a = 5$ and $b = 3$. The probability can then be computed:

$$\begin{aligned} P(S_{12} = 5|S_4 = 3) &= \binom{8}{5} 0.6^5 0.4^3 \\ &= 0.2787. \end{aligned} \tag{9}$$

3. Consider the continuous-time random walk on the graph which is a pentagon with one diagonal. Exit rate along each edge is equal to 1.

a) Write the generating matrix for this Markov chain.

Solution

For this problem, we will arbitrarily assign the diagonal to go from state 2 to state 5. The generating matrix for this MC can be found by inspection to be

$$\begin{bmatrix} -2 & 1 & 0 & 0 & 1 \\ 1 & -3 & 1 & 0 & 1 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \\ 1 & 1 & 0 & 1 & -3 \end{bmatrix} \tag{10}$$

.

b) Find its stationary distribution.

Solution

The stationary distribution for A can be found by using the equations $\pi A = 0$ where $\pi = \begin{bmatrix} \pi_1 & \pi_2 & \pi_3 & \pi_4 & \pi_5 \end{bmatrix}$ and $\pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5 = 1$.

$$\begin{bmatrix} \pi_1 & \pi_2 & \pi_3 & \pi_4 & \pi_5 \end{bmatrix} \begin{bmatrix} -2 & 1 & 0 & 0 & 1 \\ 1 & -3 & 1 & 0 & 1 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \\ 1 & 1 & 0 & 1 & -3 \end{bmatrix} = 0$$

$$\left\{ \begin{matrix} -2\pi_1 + \pi_2 + \pi_5 & = 0 \\ \pi_1 - 3\pi_2 + \pi_4 + \pi_5 & = 0 \\ \pi_2 - 2\pi_3 + \pi_4 & = 0 \\ \pi_3 - 2\pi_4 + \pi_5 & = 0 \\ \pi_1 + \pi_2 + \pi_4 - 3pi_5 & = 0 \\ \pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5 & = 1 \end{matrix} \right\} \tag{11}$$

$$\left\{ \begin{matrix} \pi_1 = 0.2 \\ \pi_2 = 0.2 \\ \pi_3 = 0.2 \\ \pi_4 = 0.2 \\ \pi_5 = 0.2 \end{matrix} \right\}$$

The corresponding stationary distribution is $\begin{bmatrix} 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \end{bmatrix}$.

c) Find eigenvalues and eigenvectors of this matrix (using online matrix calculator).

Solution

Using the provided online calculator, the eigenvalues and corresponding row eigenvectors were found to be

$$
\begin{aligned}
\lambda_1 &= 0, v_1 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \end{bmatrix} \\
\lambda_2 &= \frac{\sqrt{5}-7}{2} \approx -2.38, v_2 \approx \begin{bmatrix} 0 & -1 & -1.62 & 1.62 & 1 \end{bmatrix} \\
\lambda_3 &= -\frac{\sqrt{5}-7}{2} \approx -4.618, v_3 \approx \begin{bmatrix} 0 & -1 & 0.62 & -0.62 & 1 \end{bmatrix} \\
\lambda_4 &= \frac{\sqrt{5}-5}{2} \approx -1.318, v_4 \approx \begin{bmatrix} 3.24 & 1 & -2.62 & -2.62 & 1 \end{bmatrix} \\
\lambda_5 &= -\frac{\sqrt{5}-5}{2} \approx -3.618, v_5 \approx \begin{bmatrix} 1.24 & -1 & 0.38 & 0.38 & -1 \end{bmatrix}
\end{aligned} \tag{12}
$$

d) Find the rate of convergence to this stationary distribution.

Solution

The rate of convergence for a CTMC is e^{at}, where a is the eigenvalue closest to zero for all $\lambda_i \neq 0$. In this case, $\lambda_4 \approx -1.3819$ is the closest eigenvalue to zero, therefore the rate of convergence for this Markov chain is $e^{-1.38t}$.

e) Given that initial distribution was uniform over the five vertices, find the distribution at time t.

Solution

The distribution at time t can be found using the relationship $p'(t) = p(t)A$. By expanding this for our 5 states, the expression and it's derivative becomes

$$
\begin{aligned}
p(t) &= c_1(t)v_1 + c_2(t)v_2 + c_3(t)v_3 + c_4(t)v_4 + c_5(t)v_5 \\
p'(t) &= c_1'(t)v_1 + c_2'(t)v_2 + c_3'(t)v_3 + c_4'(t)v_4 + c_5'(t)v_5.
\end{aligned} \tag{13}
$$

Since $p'(t) = p(t)A$, this expression can be rewritten as

$$
p(t)A = c_1(t)v_1A + c_2(t)v_2A + c_3(t)v_3A + c_4(t)v_4A + c_5(t)v_5A. \tag{14}
$$

Then, using the relationship $v_kA = \lambda_k v_k$, we can see

$$
\begin{aligned}
v_1A &= \lambda_1 v_1 = 0v_1 \\
v_2A &= \lambda_2 v_2 = -2.38v_2 \\
v_3A &= \lambda_3 v_3 = -4.618v_3 \\
v_4A &= \lambda_4 v_4 = -1.318v_4 \\
v_5A &= \lambda_5 v_5 = -3.618v_5
\end{aligned} \tag{15}
$$

Substituting these values into equation 14, we get the expression

$$p'(t) = 0 - 2.38c_2(t)v_2 - 4.618c_3(t)v_3 - 1.318c_4(t)v_4 - 3.618c_5(t)v_5. \tag{16}$$

By comparing this expression to 13, we can see that $c_1'(t) = 0$, therefore $c_1(t)$ is a constant. Similar expressions can be found for the other variables: $c_2'(t) = -2.38c_2(t)$, $c_3'(t) = -4.618c_3(t)$, $c_4'(t) = -1.318c_4(t)$ and $c_5'(t) - 3.618c_5(t)$. Since these derivatives are not equal to zero, c_2, c_3, c_4 and c_5 are not constants, but can be easily solved for as first order ODEs (though the steps below show it is not necessary for this problem).

We can plug in our known initial uniform distribution of $p(0) = \begin{bmatrix}0.2 & 0.2 & 0.2 & 0.2 & 0.2\end{bmatrix}$ and eigenvectors to this expression

$$\begin{aligned}\begin{bmatrix}0.2 & 0.2 & 0.2 & 0.2 & 0.2\end{bmatrix} &= c_1(0)\begin{bmatrix}1 & 1 & 1 & 1 & 1\end{bmatrix} \\ &+ c_2(0)\begin{bmatrix}0 & -1 & -1.62 & 1.62 & 1\end{bmatrix} \\ &- c_3(0)\begin{bmatrix}0 & -1 & 0.62 & -0.62 & 1\end{bmatrix} \\ &+ c_4(0)\begin{bmatrix}3.24 & 1 & -2.62 & -2.62 & 1\end{bmatrix} \\ &- c_5(0)\begin{bmatrix}1.24 & -1 & 0.38 & 0.38 & -1\end{bmatrix}.\end{aligned} \tag{17}$$

$$=> \left\{\begin{matrix} c_1(0) = 1/5 \\ c_2(0) = 0 \\ c_3(0) = 0 \\ c_4(0) = 0 \\ c_5(0) = 0 \end{matrix}\right\}$$

As mentioned above, since $c_2(0), c_3(0), c_4(0)$ and $c_5(0)$ are equal to 0, we do not need to solve any ODEs. The distribution at time t is then given as $p(t) = c_1 v_1 = 1/5\begin{bmatrix}1 & 1 & 1 & 1 & 1\end{bmatrix} = \begin{bmatrix}0.2 & 0.2 & 0.2 & 0.2 & 0.2\end{bmatrix}$. The distribution at any time t will converge to the stationary distribution for given initial conditions that are uniformly distributed.

Code

1. Make 10000 simulations of the arithmetic and geometric random walks from Theory 1 for $T = 100$ steps, and compute all empirical quantities in Theory 1. Plot the histograms. Compare with theoretical results. Plot 5 paths for the arithmetic random walk on the same plot. Next, plot 5 paths the geometric random walk on a different plot, chosen as follows: rank by final value and choose paths corresponding to 10%, 30%, 50%, 70%, 90%.

Solution

part a

```
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

T = 100
mu = 0
scale = 1/np.sqrt(2) #for var=1
N = 10000
Sn = []
Mn = []
np.random.seed(10)
for i in range(0,N):
    Z= np.random.laplace(mu,scale,T+1) # 100 steps starting from 0
    Z[0]= 0 #S_0 - initial position zero
    S_n = np.cumsum(Z)
    Sn.append(S_n)
    Mn.append(np.exp(S_n))

print(np.mean(Sn, axis=0)[-1])
print(np.median(Sn,axis=0)[-1])
print(np.std(Sn, axis=0)[-1])

>> 0.045354445921 69611
>> -0.006378167801358914
>> 9.982679411194056

print(np.mean(Mn, axis=0)[-1])
print(np.median(Mn, axis=0)[-1])
print(np.std(Mn, axis=0)[-1])

>> 267471432458.6772
>> 0.9936426380737612
>> 24141980484724.37

arithmetic_position = [Sn[i][-1] for i in range(N)]
sns.histplot(arithmetic_position, bins=50, stat='count', kde=True)
plt.title('Arithmetic final position across N=10000 sims')
```

```
plt.xlabel('position')

geometric_position = [Mn[i][-1] for i in range(N)]
sns.histplot(geometric_position, bins=50, stat='count', kde=True)
plt.title('Geometric final position across N=10000 sims')
plt.ylabel('position')
```

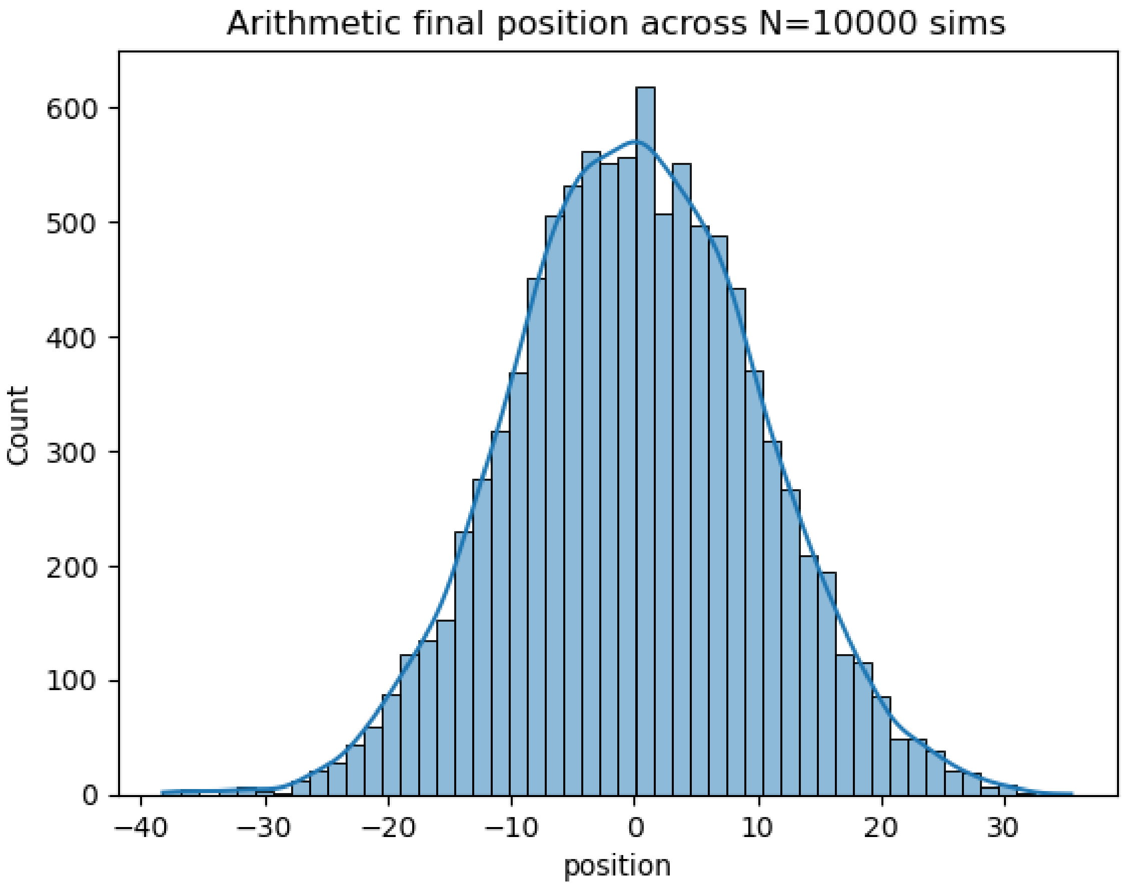

Figure 1: Distribution of 10000 different simulated arithmetic random walks with laplace random variables. As expected, the shape is approximately Laplacian with $\mu = 0$ and $\sigma = 10$.

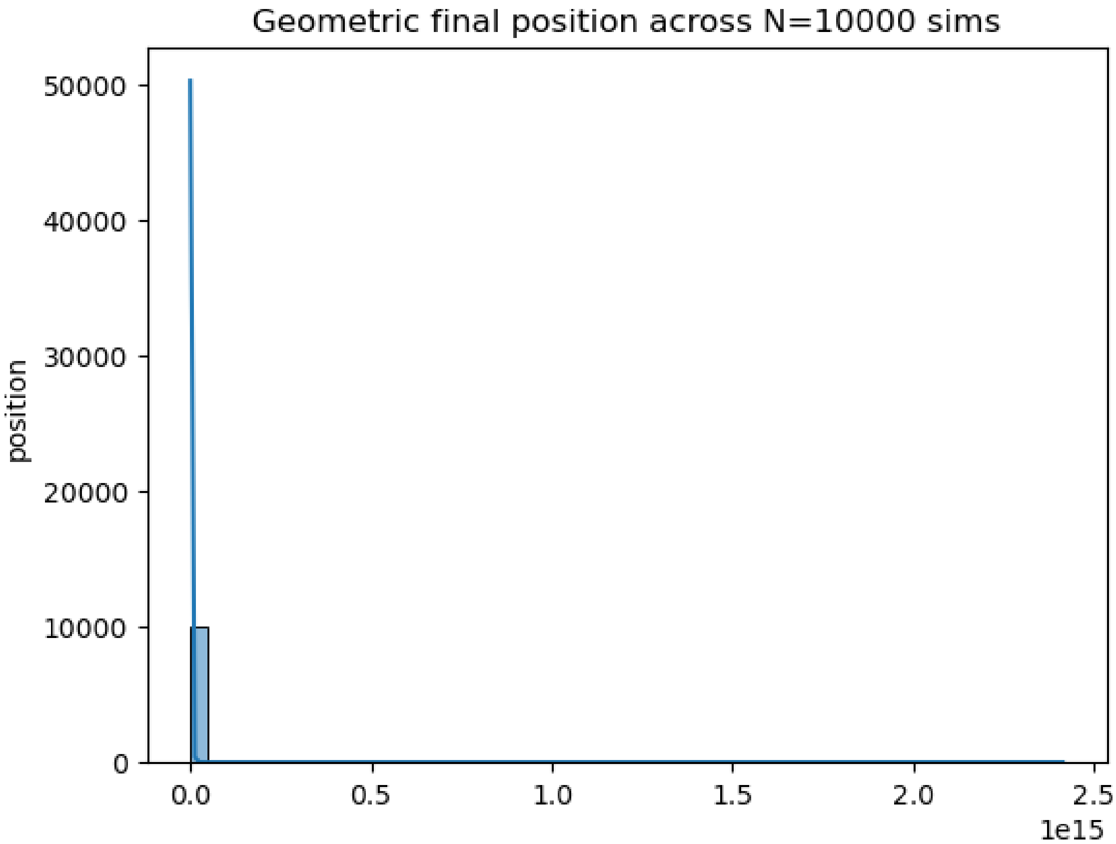

Figure 2: Distribution of 10000 different simulated geometric random walks with laplace random variables. Note μ is still near 0, but σ does not converge.

2. Simulate $T = 12$ steps of the simple random walk from Theory 2. Compute the empirical probabilities from parts A and B, and compare with theoretical results.

part b

```
np.random.seed(1)
rnd_choice = np.random.choice(np.arange(N), 5)
for i in rnd_choice:
    plt.plot(Sn[i])

plt.title('Arithmetic Paths')
plt.xlabel('steps')
plt.ylabel('position')

Sn.sort(key=lambda x: x[-1])
Mn.sort(key=lambda x: x[-1])

ten_percent = []
thirty_percent = []
fifty_percent = []
seventy_percent = []
ninety_percent = []

for i in range(N):
    if (Sn[i][-1] > 0.09) & (Sn[i][-1] < 0.11):
        ten_percent.append(Mn[i])
    if (Sn[i][-1] > 0.29) & (Sn[i][-1] < 0.31):
        thirty_percent.append(Mn[i])
    if (Sn[i][-1] > 0.49) & (Sn[i][-1] < 0.51):
        fifty_percent.append(Mn[i])
    if (Sn[i][-1] > 0.69) & (Sn[i][-1] < 0.71):
        seventy_percent.append(Mn[i])
    if (Sn[i][-1] > 0.89) & (Sn[i][-1] < 0.91):
        ninety_percent.append(Mn[i])

plt.plot(ten_percent[0], label='10% return')
plt.plot(thirty_percent[0], label='30% return')
plt.plot(fifty_percent[0], label='50% return')
plt.plot(seventy_percent[0], label='70% return')
plt.plot(ninety_percent[0], label='90% return')
plt.legend()
plt.xlabel('time')
plt.ylabel('position')
plt.title('Geometric Paths')
plt.yscale('log')
```

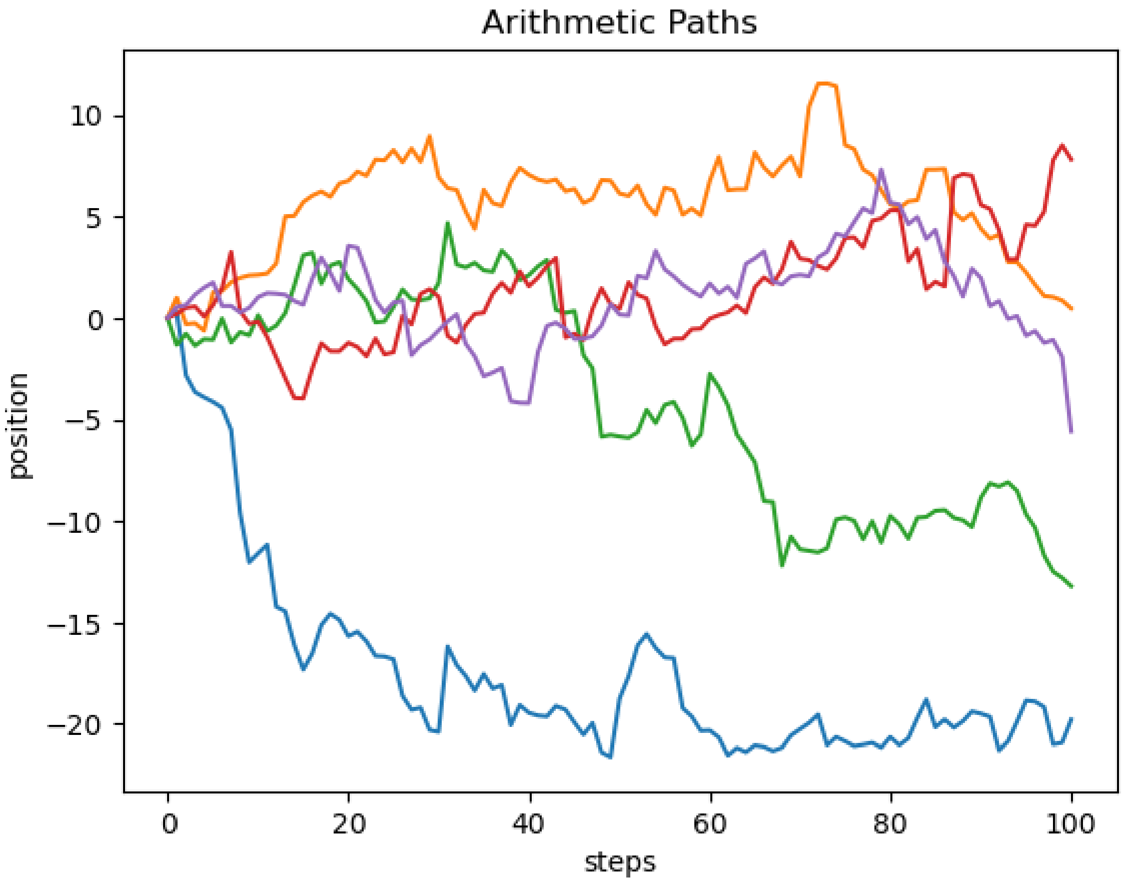

Figure 3: Five randomly chosen arithmetic random walks.

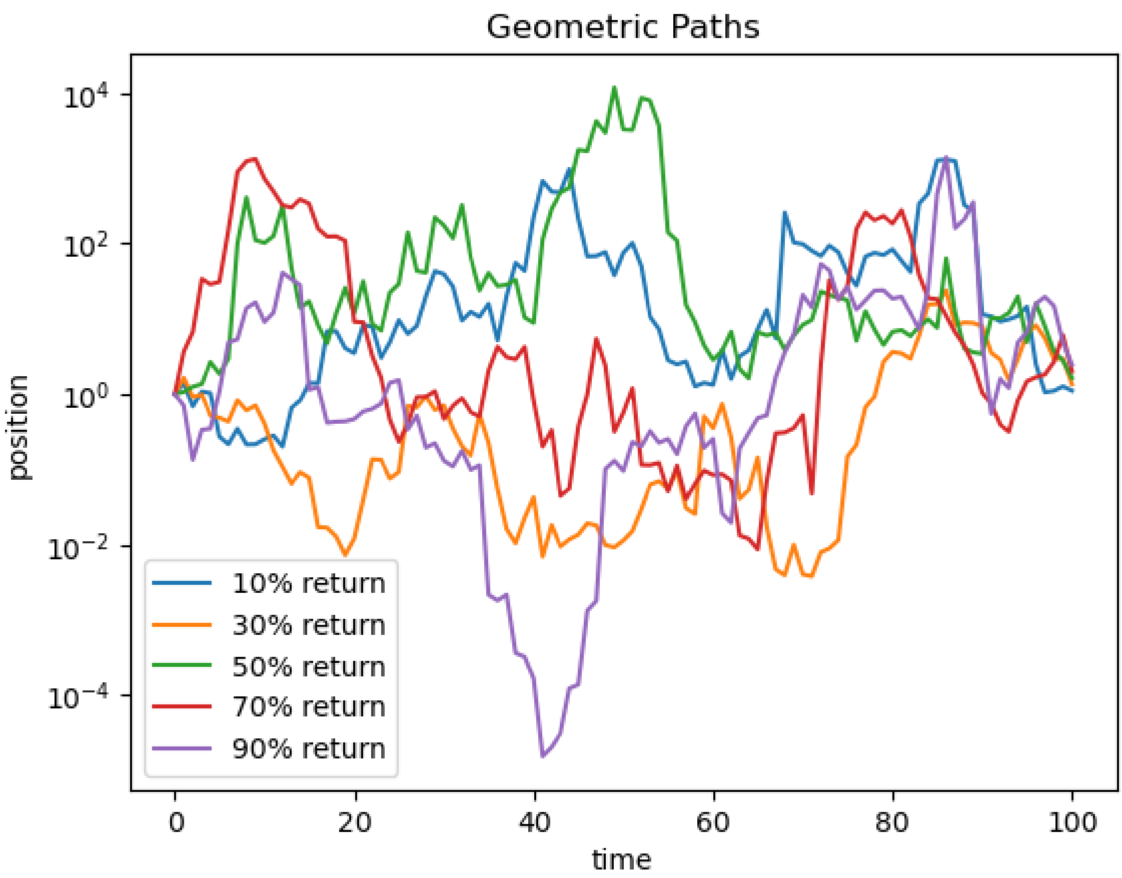

Figure 4: Five geometric random walks, chosen based on final value relative to starting value of 0. Plotted on a semi-log scale.

3. Simulate $T = 12$ steps of the simple random walk from Theory 2. Compute the empirical probabilities from parts A and B, and compare with theoretical results.
Solution

part a

```
T = 12
N = 1000
p=0.6
q=0.4
Sn = []
Mn = []
np.random.seed(1)
for i in range(0,N):
    Z= np.random.choice(a=[-1,1],p=[q,p],size=T+1)
    Z[0]= -1 #S_0 - start position
    Sn.append(np.cumsum(Z))

count = 0
for i in range(N):
    if Sn[i][-1] == 5:
        count+=1

theoretical_prob= 0.1418939597
print('P(S_12=5|S_1=-1)=', count/N)
print('Percent difference between theory and emperical:',
 ↪  (((count/N)-theoretical_prob)/theoretical_prob))

>> P(S_12=5|S_1=-1)= 0.143
>> Percent difference between theory and emperical: 0.0077948370905882888
```

part b

```
T = 8
N = 1000
p=0.6
q=0.4
Sn = []
Mn = []
np.random.seed(1)
for i in range(0,N):
    Z= np.random.choice(a=[-1,1],p=[q,p],size=T+1)
    Z[0]= 3 #S_0 - start position
    Sn.append(np.cumsum(Z))

count = 0
for i in range(N):
    if Sn[i][-1] == 5:
        count+=1

theoretical_prob= 0.27869184
print('P(S_12=5|S_4=3)',count/N)
```

```
print('Difference between theory and emperical:',
↪  (((count/N)-theoretical_prob)/theoretical_prob))

>> P(S_12=5|S_4=3) 0.281
>> Percent difference between theory and emperical: 0.008282122648442105
```