HW 3 M362M, Fall 2024

Wei Xuan Lee (wl22963)

2024-09-19

Problem 1

1.1

$\{2X_n\}_{n\in\mathbb{N}_0}$ **<u>is NOT</u> a simple random walk.** For any $i \in \mathbb{N}_0$, let $\delta'_i = 2X_i - 2X_{i-1} = 2(X_i - X_{i-1}) = 2\delta_i$. Note that δ'_i does not take on any value other than -2, 2 since δ_i does not take on any value other than -1, 1. Therefore δ'_i does not follow a coin-toss distribution as it does not take on the values $-1, 1$.

1.2

$\{X_n^2\}_{n\in\mathbb{N}_0}$ **<u>is NOT</u> a simple random walk.** For any $i \in \mathbb{N}_0$, let $\delta'_i = X_i^2 - X_{i-1}^2 = \delta_i \cdot (X_i + X_{i-1})$. Note that δ'_3 can take on the value 1 when $X_2 = 0, X_2 = 1$, the value 3 when $X_2 = 1, X_3 = 2$, the value 5 when $X_1 = 2, X_2 = 3$, and many other values too. Therefore it is not possible for all δ'_i, $i = 1, 2, \cdots$ to have a coin-toss distribution, which only takes on two values, namely, -1 or 1.

1.3

$\{-X_n\}_{n\in\mathbb{N}_0}$ **<u>IS</u> a simple random walk.** First, we have $-X_0 = -0 = 0$ since $X_0 = 0$. Next, for any $i \in \mathbb{N}_0$, let $\delta'_i = -X_i - (-X_{i-1}) = -(X_i - X_{i-1}) = -\delta_i$. For $i \neq j$, δ_i, δ_j are independent, therefore their constant multiples δ'_i, δ'_j are independent too. We have $\Pr[\delta'_i = 1] = \Pr[-\delta_i = 1] = \Pr[\delta_i = -1] = \frac{1}{2}$. Similarly $\Pr[\delta'_i = -1] = \frac{1}{2}$. δ'_i does not take on any value other than 1, -1 since δ_i does not take on any value other than -1, 1. Thus all $\delta'_i, i \in \mathbb{N}_0$ have a coin-toss distribution.

1.4

$\{Y_n\}_{n\in\mathbb{N}_0}$ **<u>IS</u> a simple random walk.** First, we have $Y_0 = X_5 - X_5 = 0$. Next, for any $i \in \mathbb{N}_0$, let $\delta'_i = Y_i - (Y_{i-1}) = (X_{i+5} - X_5) - (X_{i+4} - X_5) = X_{i+5} - X_{i+4} = \delta_{i+5}$. For $i \neq j$, δ_i, δ_j are independent, therefore their constant multiples δ'_i, δ'_j are independent too. We have $\Pr[\delta'_i = 1] = \Pr[-\delta_i = 1] = \Pr[\delta_i = -1] = \frac{1}{2}$. Similarly $\Pr[\delta'_i = -1] = \frac{1}{2}$. δ'_i does not take on any value other than 1, -1 since δ_i does not take on any value other than -1, 1. Thus all $\delta'_i, i \in \mathbb{N}_0$ have a coin-toss distribution.

Problem 2

2.1

$$
\begin{aligned}
&\Pr[X_{2n} = 0] \\
=&\Pr\left[\sum_{i=1}^{2n}(X_i - X_{i-1}) = 0\right] \\
=&\Pr\left[\sum_{i=1}^{2n}\delta_i = 0\right] \\
=&\sum_{\substack{(k_1,\cdots,k_{2n})\in\{-1,1\}^{2n},\\ k_1+\cdots k_{2n}=0}} \Pr[\delta_1 = k_1, \cdots, \delta_{2n} = k_{2n}] \\
=&\sum_{\substack{(k_1,\cdots,k_{2n})\in\{-1,1\}^{2n},\\ k_1+\cdots k_{2n}=0}} \left(\prod_{i=1}^{2n}\Pr[\delta_i = k_i]\right) && (\text{Independence of } \{\delta_i\}_{i=1}^{2n}) \\
=&\sum_{\substack{(k_1,\cdots,k_{2n})\in\{-1,1\}^{2n},\\ k_1+\cdots k_{2n}=0}} 2^{-2n} && (\Pr[\delta_i = j] = 2^{-1}, \forall (i,j) \in \{1,\cdots,2n\}\times\{-1,1\}) \\
=&\binom{2n}{n}\cdot 2^{-2n} && (\text{Choose } n \text{ of } k_1,\cdots,k_{2n} \text{ be 1 and the rest be } -1)
\end{aligned}
$$

```r
# Here we pick n = 8, therefore 2n (num_steps) = 16
n = 8
num_steps = 2 * n
num_sim = 10000

simulate_path = function(num_steps) {
  steps = sample(c(-1,1), num_steps, replace=TRUE, prob=c(.5, .5))
  return (c(0, cumsum(steps)))
}

path_ends_zero = function(num_steps) {
  path = simulate_path(num_steps)
  # Returns TRUE if path ends at 0
  return (path[num_steps + 1] == 0)
}

# Calculate probability (Simulation)
sim_prob = mean(replicate(num_sim, path_ends_zero(num_steps)))

# Calculate probability (Theoretical)
the_prob = choose(num_steps, n) * (2 ** (-num_steps))

cat(" We calculate P[X_16 = 0]", "\n", "\n",
      "Probability (Theoretical) :", round(the_prob, 4), "\n",
      "Probability (Simulation)  :", round(sim_prob, 4))
##  We calculate P[X_16 = 0]
##
##  Probability (Theoretical) : 0.1964
##  Probability (Simulation)  : 0.1963
```

2.2

$$
\begin{aligned}
&\Pr[X_n = X_{2n}] \\
=&\Pr[X_{2n} - X_n = 0] \\
=&\Pr\left[\sum_{i=n+1}^{2n} \delta_i = 0\right] \\
=&\begin{cases} \binom{2k}{k} \cdot 2^{-2k} & \text{, if } n = 2k, k \in \mathbb{N}_0 \\ 0 & \text{, otherwise}^\dagger. \end{cases}
\end{aligned}
$$

† The sum of an odd number of -1's and 1's always gives an odd number ($\neq 0$).

```r
# Here we pick n = 8, therefore 2n (num_steps) = 16
n = 8
num_sim = 10000

simulate_path = function(num_steps) {
  steps = sample(c(-1,1), num_steps, replace=TRUE, prob=c(.5, .5))
  return (c(0, cumsum(steps)))
}

mid_equal_end = function(n) {
  # IMPORTANT: num_steps = 2*n
  path = simulate_path(2*n)

  # Returns TRUE if
  # Midpoint position = Endpoint Position
  # Note: If n is an odd number, we will always get FALSE
  return (path[n + 1] == path[2*n + 1])
}

# Calculate probability (Simulation)
sim_prob = mean(replicate(num_sim, mid_equal_end(n)))

# Calculate probability (Theoretical)
the_prob = choose(n, n/2) * (2 ** (-n))

cat(" We calculate P[X_8 = X_16]", "\n", "\n",
      "Probability (Theoretical) :", round(the_prob, 4), "\n",
      "Probability (Simulation)  :", round(sim_prob, 4))
##  We calculate P[X_8 = X_16]
##
##  Probability (Theoretical) : 0.2734
##  Probability (Simulation)  : 0.284
```

2.3

We list a table displaying all (a total of 8) possible values of X_1, X_2, X_3 :

X_1	X_2	X_3
1	2	3
1	2	1
1	0	1
1	0	-1
-1	0	1
-1	0	-1
-1	-2	-1
-1	-2	-3

Table 1: Table of all possible values of X_1, X_2, X_3.

Through observation, we see that

$$\begin{aligned}
&\Pr[|X_1 X_2 X_3| = 2] \\
=&\Pr[X_1 = 1, X_2 = 2, X_3 = 1] + \Pr[X_1 = -1, X_2 = -2, X_3 = -1] \\
=&\ \frac{1}{8} + \frac{1}{8} && \text{(Each possible sequence of } (X_1, X_2, X_3) \text{ is equally likely)} \\
=&\ \frac{1}{4}
\end{aligned}$$

```r
# Here num_steps = 16
num_steps = 16
num_sim = 10000

simulate_path = function(num_steps) {
  steps = sample(c(-1,1), num_steps, replace=TRUE, prob=c(.5, .5))
  return (c(0, cumsum(steps)))
}

abs_x1x2x3_equal_2 = function(num_steps) {
  path = simulate_path(num_steps)

  # Returns TRUE if |X_1 X_2 X_3| = 2
  x1 = path[2]
  x2 = path[3]
  x3 = path[4]
  return (abs(x1*x2*x3) == 2)
}

# Calculate probability (Simulation)
sim_prob = mean(replicate(num_sim, abs_x1x2x3_equal_2(num_steps)))

# Calculate probability (Theoretical)
the_prob = 0.25
```

```
cat(" We calculate P[|X_1 X_2 X_3| = 2]", "\n", "\n",
      "Probability (Theoretical) :", round(the_prob, 4), "\n",
      "Probability (Simulation)  :", round(sim_prob, 4))
##  We calculate P[|X_1 X_2 X_3| = 2]
##
##  Probability (Theoretical) : 0.25
##  Probability (Simulation)  : 0.2433
```

2.4

$$
\begin{aligned}
&\Pr[X_7 + X_12 = X_1 + X_{16}] \\
=&\Pr[X_{16} - X_{12} = X_7 - X_1] \\
=&\Pr[\delta_{16} + \delta_{15} + \delta_{14} + \delta_{13} + \delta'_7 + \delta'_6 + \delta'_5 + \delta'_4 + \delta'_3 + \delta'_2 = 0] && (\delta'_i = -\delta_i) \\
=&\binom{10}{5} \cdot 2^{-10} \\
=&\frac{63}{256} \approx 0.2461 \text{ (4 d.p.)}
\end{aligned}
$$

```
# Here num_steps = 16
num_steps = 16
num_sim = 10000

simulate_path = function(num_steps) {
  steps = sample(c(-1,1), num_steps, replace=TRUE, prob=c(.5, .5))
  return (c(0, cumsum(steps)))
}

x7_plus_x12_equal_x1_plus_x_16 = function(num_steps) {
  path = simulate_path(num_steps)

  # Returns TRUE if X_7 + X_12 = X_1 + X_16
  x1 = path[2]
  x7 = path[8]
  x12 = path[13]
  x16 = path[17]
  return (x7 + x12 == x1 + x16)
}

# Calculate probability (Simulation)
sim_prob = mean(replicate(num_sim, x7_plus_x12_equal_x1_plus_x_16(num_steps)))

# Calculate probability (Theoretical)
the_prob = 63/256

cat(" We calculate P[X_7 + X_12 = X_1 + X_16]", "\n", "\n",
      "Probability (Theoretical) :", round(the_prob, 4), "\n",
      "Probability (Simulation)  :", round(sim_prob, 4))
##  We calculate P[X_7 + X_12 = X_1 + X_16]
##
##  Probability (Theoretical) : 0.2461
##  Probability (Simulation)  : 0.2445
```

Problem 3

We shall first try to count the number of paths such that $X_i > 0$ for $i = 1, \cdots, n$. Note that we must have $X_1 = 1$. Let $X_n = k > 0$ (n, k have the same parity), notice that there is a bijection between the paths from $X_1 = 1$ to $X_n = k$ that **DO** visit 0 (suppose $X_{j_0} = 0$ and the path never visits 0 before j_0-th step) and all the paths from $X_1 = -1$ to $X_n = k$. The bijection is as follows:

$$(X_1, X_2, \cdots, X_{j_0}, X_{j_0+1}, \cdots, X_n) \mapsto (-X_1, -X_2, \cdots, -X_{j_0}, X_{j_0+1}, \cdots, X_n).$$

We know that the number of paths from $X_1 = -1$ to $X_n = k$ is $\binom{n-1}{\frac{n+k}{2}}$, due to the bijection, is also the number of paths from $X_1 = 1$ to $X_n = k$ that **DO** visit 0. Since we are looking for those that do **NOT** visit 0, we subtract it from the number paths from $X_1 = 1$ to $X_n = k$, which gives us (Note: k > 0)

$$\begin{aligned}
&\#\{\text{From } X_0 = 0 \text{ to } X_n = k, \text{does not visit } 0\} \\
=&\#\{\text{From } X_1 = 1 \text{ to } X_n = k, \text{does not visit } 0\} \\
=&\#\{\text{From } X_1 = 1 \text{ to } X_n = k\} - \#\{\text{From } X_1 = 1 \text{ to } X_n = k, \text{visits } 0\} \\
=&\binom{n-1}{\frac{n+k}{2}-1} - \binom{n-1}{\frac{n+k}{2}} \\
=&\frac{(n-1)!}{(n-\frac{n+k}{2})!\,(\frac{n+k}{2})!} \cdot \left[\frac{n+k}{2} - \left(n - \frac{n+k}{2}\right)\right] \\
=&\frac{k}{n} \cdot \binom{n}{\frac{n+k}{2}}.
\end{aligned}$$

Thus, we get that for $k > 0$,

$$\begin{aligned}
&\Pr[\text{Does not visit } 0 \mid \text{From } X_0 = 0 \text{ to } X_n = k] \\
=&\frac{\#\{\text{From } X_0 = 0 \text{ to } X_n = k, \text{does not visit } 0\}}{\#\{\text{From } X_0 = 0 \text{ to } X_n = k\}} \\
=&\frac{\frac{k}{n} \cdot \binom{n}{\frac{n+k}{2}}}{\binom{n}{\frac{n+k}{2}}} \\
=&\frac{k}{n}.
\end{aligned}$$

With a similar reasoning, we get that for $k < 0$,

$$\Pr[\text{Does not visit } 0 \mid \text{From } X_0 = 0 \text{ to } X_n = k] = \frac{|k|}{n}.$$

Therefore we have that, if $n = 2m - 1, m = 1, 2, \cdots$ (odd number)

$$
\begin{aligned}
&\Pr[\text{Starts from 0, does not visit 0 in the next } n \text{ steps}] \\
=&\sum_{i=-m-1}^{m} \Pr[\text{From } X_0 = 0 \text{ to } X_n = 2i - 1,\ \text{does not visit 0}] \\
=&\sum_{i=-m-1}^{m} \Pr[\text{Does not visit 0} \mid \text{From } X_0 = 0 \text{ to } X_n = 2i + 1] \\
&\qquad \times \Pr[\text{From } X_0 = 0 \text{ to } X_n = 2i + 1] \\
=&2 \times \sum_{i=1}^{m} \frac{2i - 1}{n} \times \frac{\binom{n}{\frac{n-1}{2}+i}}{2^n} \\
=&\frac{1}{n \cdot 2^{n-1}} \sum_{i=1}^{m} \left[(2i - 1) \times \binom{n}{\frac{n-1}{2} + i} \right]
\end{aligned}
$$

On the other hand, if $n = 2m, m = 1, 2 \cdots$ (even number)

$$
\begin{aligned}
&\Pr[\text{Starts from 0, does not visit 0 in the next } n \text{ steps}] \\
=&\sum_{i=-m, i \neq 0}^{m} \Pr[\text{From } X_0 = 0 \text{ to } X_n = 2i,\ \text{does not visit 0}] \\
=&\sum_{i=-m, i \neq 0}^{m} \Pr[\text{Does not visit 0} \mid \text{From } X_0 = 0 \text{ to } X_n = 2i] \\
&\qquad \times \Pr[\text{From } X_0 = 0 \text{ to } X_n = 2i] \\
=&2 \times \sum_{i=1}^{m} \frac{2i}{n} \times \frac{\binom{n}{\frac{n}{2}+i}}{2^n} \\
=&\frac{1}{n \cdot 2^{n-2}} \sum_{i=1}^{m} \left[i \times \binom{n}{\frac{n}{2} + i} \right]
\end{aligned}
$$

(R Code on Next Page)

```r
# Here num_steps = 10
num_steps = 10
num_sim = 10000

# For this problem we do not start with 0
simulate_path = function(num_steps) {
  steps = sample(c(-1,1), num_steps, replace=TRUE, prob=c(.5, .5))
  # Note that we do not start with 0
  return (cumsum(steps))
}

skip_zero = function(num_steps) {
  path = simulate_path(num_steps)

  # If does not visit 0, then
  # 'check' should be all FALSE
  check = (path == 0)

  # Returns TRUE if does not visit 0
  return(sum(check) == 0)
}

# Calculate probability (Simulation)
sim_prob = mean(replicate(num_sim, skip_zero(num_steps)))

# Calculate probability (Theoretical)
the_sum = 0

# If num_steps is even
if ((num_steps %% 2) == 0) {
  for (i in 1:(num_steps/2)) {
    the_sum = the_sum + (i * choose(num_steps, (num_steps/2) + i))
  }

  the_prob = (the_sum)/(num_steps * (2 ** (num_steps - 2)))
} else {
  # If num_steps is odd
  for (i in 1:((num_steps+1)/2)) {
    the_sum = the_sum + ((2*i - 1) * choose(num_steps, (num_steps-1)/2 + i))
  }

  the_prob = (the_sum)/(num_steps * (2 ** (num_steps - 1)))
}

cat(" We calculate P[Does not visit 0, n=10]", "\n", "\n",
      "Probability (Theoretical) :", round(the_prob, 4), "\n",
      "Probability (Simulation)  :", round(sim_prob, 4))
##  We calculate P[Does not visit 0, n=10]
##
##  Probability (Theoretical) : 0.2461
##  Probability (Simulation)  : 0.2484
```

Problem 4

4.1

Note that

$$
\begin{aligned}
\Pr[X=i] &= \sum_{j=1}^{3} \Pr[X=i, Y=j], \text{ for } i=1,2,3 \\
\Pr[Y=j] &= \sum_{i=1}^{3} \Pr[X=i, Y=j], \text{ for } j=1,2,3
\end{aligned}
$$

Thus,

$$
\begin{array}{ll}
\Pr[X=1]=0.4 & \Pr[Y=1]=0.40 \\
\Pr[X=2]=0.3 & \Pr[Y=2]=0.45 \\
\Pr[X=3]=0.3 & \Pr[Y=3]=0.15
\end{array}
$$

Therefore

$$
\begin{aligned}
\mathbb{E}[X] =& 1\times 0.4 + 2\times 0.3 + 3\times 0.3 \\
=& 1.9
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}[Y] =& 1\times 0.40 + 2\times 0.45 + 3\times 0.15 \\
=& 1.75
\end{aligned}
$$

4.2

First, we calculate $\mathbb{E}[XY]$.

$$
\begin{aligned}
\mathbb{E}[XY] =& 1\times 1\times 0.10 + 1\times 2\times 0.20 + 1\times 3\times 0.10 \\
& 2\times 1\times 0.15 + 2\times 2\times 0.10 + 2\times 3\times 0.05 \\
& 3\times 1\times 0.15 + 3\times 2\times 0.15 + 3\times 3\times 0 \\
=& 3.15.
\end{aligned}
$$

Therefore we have

$$
\begin{aligned}
\mathrm{Cov}(X,Y) =& \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \\
=& 3.15 - 1.9\times 1.75 \\
=& -0.175
\end{aligned}
$$

4.3

$$
\begin{aligned}
\mathbb{E}[X \mid Y=2] =& \sum_{i=1}^{3} i \times \Pr[X=i \mid Y=2] \\
=& \sum_{i=1}^{3} i \times \frac{\Pr[X=i, Y=2]}{\Pr[Y=2]} \\
=& 1 \times \frac{0.2}{0.45} + 2 \times \frac{0.1}{0.45} + 3 \times \frac{0.15}{0.45} \\
=& 1.8889 \text{ (4 d.p.)}
\end{aligned}
$$

4.4

```r
# Create a table for the joint distribution of X, Y
joint_pmf <- data.frame(
  X = c(1, 1, 1, 2, 2, 2, 3, 3, 3),
  Y = c(1, 2, 3, 1, 2, 3, 1, 2, 3),
  Prob = c(0.1, 0.2, 0.3, 0.15, 0.1, 0.05, 0.15, 0.15, 0)
)

# Number of draws
num_sim = 10000

joint_sample = function(num_sim) {
  # Pick num_sim numbers between 1 and the number of rows of joint_pmf
  pick = sample(1:nrow(joint_pmf), num_sim, replace=TRUE, prob=joint_pmf$Prob)

  # Return the X,Y's
  return (joint_pmf[pick,][c("X", "Y")])
}

# Sample
XY_sample = joint_sample(num_sim)

# Only keep those with Y = 2
X_Yeq2_sample = XY_sample[XY_sample$Y == 2,]

# Calculate E[X | Y =2] (Simulation)
sim_exp = mean(X_Yeq2_sample$X)

# Calculate E[X | Y =2] (Theoretical)
the_exp = 1.8889

cat(" We calculate E[X | Y = 2]", "\n", "\n",
      "Expectation (Theoretical) :", round(the_exp, 4), "\n",
      "Expectation (Simulation)  :", round(sim_exp, 4))
##  We calculate E[X | Y = 2]
##
##  Expectation (Theoretical) : 1.8889
##  Expectation (Simulation)  : 1.8931
```

End of Homework