CSE 250A HW3

Jiping Lin A15058075

1. (a) when $t=1$, $P(X_2=j|X_1=i)=[A]_{ij}$.

Suppose $P(X_{k+1}=j|X_1=i)=[A^k]_{ij}$ for some $k>1$,

then $P(X_{k+2}=j|X_1=i)=\sum_{\ell=1}^{m} P(X_{k+2}=j, X_{k+1}=\ell|X_1=i)$

$$= \sum_{\ell=1}^{m} P(X_{k+1}=\ell|X_1=i)P(X_{k+2}=j|X_{k+1}=\ell, X_1=i) \quad \text{「Product Rule」}$$

$$= \sum_{\ell=1}^{m} [A^k]_{i\ell} P(X_{k+2}=j|X_{k+1}=\ell) \quad \text{(by d.sep (1))} \quad X_1 \to X_{k+2} | X_{k+1}$$

$$= \sum_{\ell=1}^{m} [A^k]_{i\ell}[A]_{\ell j}$$

$$= [A^{k+1}]_{ij},$$

By induction, we have $P(X_{t+1}=j|X_1=i)=[A^t]_{ij}$, $\forall t \geq 1$.

(b)
```
int prob(int[][] A, int t, int j, int i){
    if (t==1) return A[i][j];
    int res = 0;
    for (int k = 0; k < A.length; k++)          (*)
        res += prob(A, t-1, k, i) * prob(A, 1, j, k);
    return res;
}
```

Since $[A^t]_{ij}=\sum_{k=1}^{m}[A^{t-1}]_{ik}[A]_{kj}$, we use recursion to compute each term in summation, each row * col is $O(m)$. prob(A, t-1, k, i) needs $O(t)$ to reduce to base case, and sum takes m steps, so it's $O(m^2 t)$.

(c) Similarly, we split the A^k to $A^{k/2}$ and $A^{k/2}$ in recursive steps:

```
int prob2(int[][] A, int t, int j, int i) {
    if (t==1) return A[i][j];
    int res = 0;
    if((t&1)==1){
        for (int k=0; k<A.length; k++){
            res += prob2(A, t-1, k, i) * prob2(A, 1, j, k);
        }
    } else {
        for(int k=0; k<A.length; k++){
            res += prob2(A, t/2, k, i)*prob2(A, t/2, j, k);
        }
    }
    return res;
}
```

base case (t=1), so it is $O(m^2)$

This is $O(m^3 \log t)$ since recursion takes $\log(t)$ but in summation we don't always have

(d) Instead of going over the entire row, we only look at nonzero entries.
We add following to function in (b):

```
Set<Integer> nonzero = new HashSet<>();
for (int k=0; k<A.length; k++)
    if (A[i][k] != 0) nonzero.add(k);
```

then change (*) to `for (int k : nonzero)`.
This is $O(smt)$ since the row * col becomes $O(s)$.

(e) $P(X_1=i \mid X_{t+1}=j) = \dfrac{P(X_{t+1}=j \mid X_1=i)\, P(X_1=i)}{P(X_{t+1}=j)}$, 「Baye's」

where $P(X_{t+1}=j) = \sum_{i=1}^{m} P(X_{t+1}=j, X_1=i)$

$= \sum_{i=1}^{m} P(X_{t+1}=j \mid X_1=i)\, P(X_1=i)$

$= \sum_{i=1}^{m} [A^t]_{ij}\, P(X_1=i)$,

So $P(X_1=i \mid X_{t+1}=j) = \dfrac{[A^t]_{ij}\, P(X_1=i)}{\sum_{k=1}^{m} [A^t]_{kj}\, P(X_1=k)}$

2. (a) $P(Y_1 \mid X_1) = \sum_x P(Y_1, X_0=x \mid X_1)$

$= \sum_x P(Y_1 \mid X_0=x, X_1)\, P(X_0=x \mid X_1)$

$= \sum_x P(Y_1 \mid X_0=x, X_1)\, P(X_0=x)$

(b) $P(Y_1) = \sum_x P(Y_1, X_1=x)$

$= \sum_x P(Y_1 \mid X_1=x)\, P(X_1=x)$

(c) $P(X_n \mid Y_1, \cdots, Y_{n-1}) = P(X_n)$

By d.sep(3), $X_n \to Y_i \mid \emptyset$ is blocked by Y_n, $\forall i$.

(d) $P(Y_n \mid X_n, Y_1, \cdots, Y_{n-1}) = \sum_x P(Y_n, X_{n-1}=x \mid X_n, Y_1, \cdots, Y_{n-1})$

$= \sum_x P(Y_n \mid X_n, Y_1, \cdots, Y_{n-1}, X_{n-1}=x)\, P(X_{n-1}=x \mid X_n, Y_1, \cdots, Y_{n-1})$

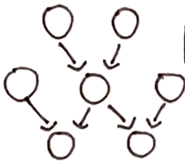
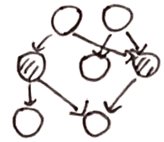
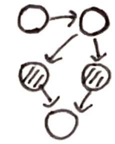
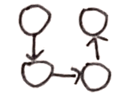
$= \sum_x P(Y_n \mid X_n, X_{n-1})\, P(X_{n-1}=x \mid Y_1, \cdots, Y_{n-1})$

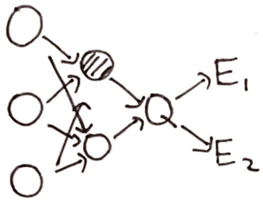
「$Y_n \to Y_1, \cdots, Y_{n-1} \mid X_n, X_{n-1}$ by d.sep(1), $X_{n-1} \to X_n \mid Y_1, \cdots, Y_{n-1}$ by d.sep(3)」

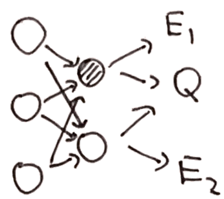
(e) $P(Y_n \mid Y_1, \cdots, Y_{n-1}) = \sum_x P(Y_n, X_n=x \mid Y_1, \cdots, Y_{n-1})$

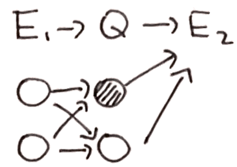
$= \sum_x P(X_n=x \mid Y_1, \cdots, Y_{n-1})\, P(Y_n \mid X_n=x, Y_1, \cdots, Y_{n-1})$

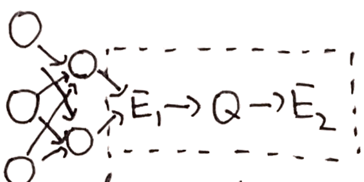
$= \sum_x \underbrace{P(X_n=x)}_{\text{by (c)}}\, \underbrace{P(Y_n \mid X_n=x, Y_1, \cdots, Y_{n-1})}_{\text{known from (d)}}$

3.  Polytree  X Polytree  x Polytree  Polytree  Polytree.

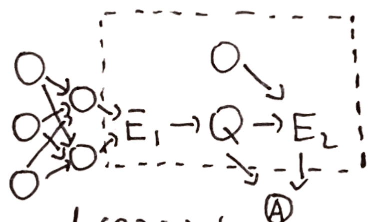
4. 







d.sep(1) by E_1



d.sep(1) by E_1.

5.

Y	$P(Y\mid X=0)$	$P(Y\mid X=1)$	$P(Z_1=1\mid Y)$	$P(Z_2=1\mid Y)$
1	$(0.25)(0.5)(0.75)=0.09375$	$(0.5)(0.75)(0.25)=0.09375$	0.9	0.1
2	$(0.75)(0.5)(0.75)=0.28125$	$(0.5)(0.75)(0.25)=0.09375$	0.8	0.2
3	$(0.25)(0.5)(0.75)=0.09375$	$(0.5)(0.25)(0.25)=0.03125$	0.7	0.3
4	$(0.25)(0.5)(0.25)=0.03125$	$(0.5)(0.75)(0.75)=0.28125$	0.6	0.4
5	$(0.75)(0.5)(0.75)=0.28125$	$(0.5)(0.25)(0.25)=0.03125$	0.5	0.5
6	$(0.75)(0.5)(0.25)=0.09375$	$(0.5)(0.75)(0.75)=0.28125$	0.4	0.6
7	$(0.25)(0.5)(0.25)=0.03125$	$(0.5)(0.25)(0.75)=0.09375$	0.3	0.7
8	$(0.75)(0.5)(0.25)=0.09375$	$(0.5)(0.25)(0.75)=0.09375$	0.2	0.8

6. (a) $\sum_z P(Z=z\mid B_1,\cdots,B_n) = \sum_z \left(\frac{1-\alpha}{1+\alpha}\right)\alpha^{|z-f(B)|}$

$$= \left(\frac{1-\alpha}{1+\alpha}\right)\sum_z \alpha^{|z-f(B)|}$$

$$= \left(\frac{1-\alpha}{1+\alpha}\right)\left[\sum_{z\le f(B)}\alpha^{f(B)-z} + \sum_{z\ge f(B)}\alpha^{z-f(B)} - \alpha^{f(B)-f(B)}\right]$$

$$= \left(\frac{1-\alpha}{1+\alpha}\right)\left[\sum_{z\le f(B)}\frac{\alpha^{f(B)}}{\alpha^{z}} + \sum_{z\ge f(B)}\frac{\alpha^{z}}{\alpha^{f(B)}} - 1\right]$$

$$= \left(\frac{1-\alpha}{1+\alpha}\right)\left[\frac{1}{1-\alpha} + \frac{1}{1-\alpha} - 1\right]$$

$$= \left(\frac{1-\alpha}{1+\alpha}\right)\left(\frac{1+\alpha}{1-\alpha}\right) = 1$$

(b)-(d) In code part

7. (a) $P(B\mid A,C,D) = \dfrac{P(D\mid B,A,C)\,P(B\mid A,C)}{P(D\mid A,C)}$ ⌜Swap B, D Baye's Rule⌟

where

$P(D\mid B,A,C) = P(D\mid B,C)$ ⌜ $D \to A \mid B, C$, d.sep (1)⌟

$P(B\mid A,C) = P(B\mid A)$ ⌜ $B \to C \mid A$, d.sep(3)⌟

$P(D\mid A,C) = \sum_b P(D, B=b\mid A,C) = \sum_b P(D\mid B=b, A, C)\,P(B=b\mid A,C) = \sum_b P(D\mid B=b, C)\,P(B=b\mid A)$

So $P(B\mid A,C,D) = \dfrac{P(D\mid B,C)\,P(B\mid A)}{\sum_b P(D\mid B=b,C)\,P(B=b\mid A)}$

(b) $P(B|A,C,D,E,F) = P(B|A,C,D)$ 「$B \to E,F|A,C,D$ by d.sep(2),
$B \to E$ blocked by C,
$B \to F$ blocked by A.」

(c) $P(B,E,F|A,C,D) = P(B|E,F,A,C,D)P(E|F,A,C,D)P(F|A,C,D)$

$$= \underbrace{P(B|A,C,D)}_{\text{by (a)}} \underbrace{P(E|C)}_{\substack{E \to A,F,D|C \\ \text{by d.sep(2)}}} \underbrace{P(F|A)}_{\substack{F \to C,D|A \\ \text{by d.sep(2)}}}.$$

8. (a) $$P(Q=q|E=e) \approx \frac{\sum_{t=1}^{T} I(q,q_t)P(E=e|Y=y_t, Z=z_t)}{\sum_{t=1}^{T} P(E=e|Y=y_t, Z=z_t)}$$

where y_t is from $P(Y|X=x_t, Q=q_t)$, $x_t \sim P(X)$,
z_t is from $P(Z|Q=q_t)$ $q_t \sim P(Q)$.

(b) $$P(Q_1=q_1, Q_2=q_2) \approx \frac{\sum_{t=1}^{T} I(q_1,q_{1t})I(q_2,q_{2t})P(E_1=e_1|Q_1=q_{1t}, X=x_t)P(E_2=e_2|E_1=e_1, Z=z_t)}{\sum_{t=1}^{T} P(E_1=e_1|Q_1=q_{1t}, X=x_t)P(E_2=e_2|E_1=e_1, Z=z_t)}$$

where q_{1t} is from $P(Q_1)$, x_t is from $P(X)$, y_t is from $P(Y)$
z_t is from $P(Z|X=x_t, Y=y_t)$

Untitled

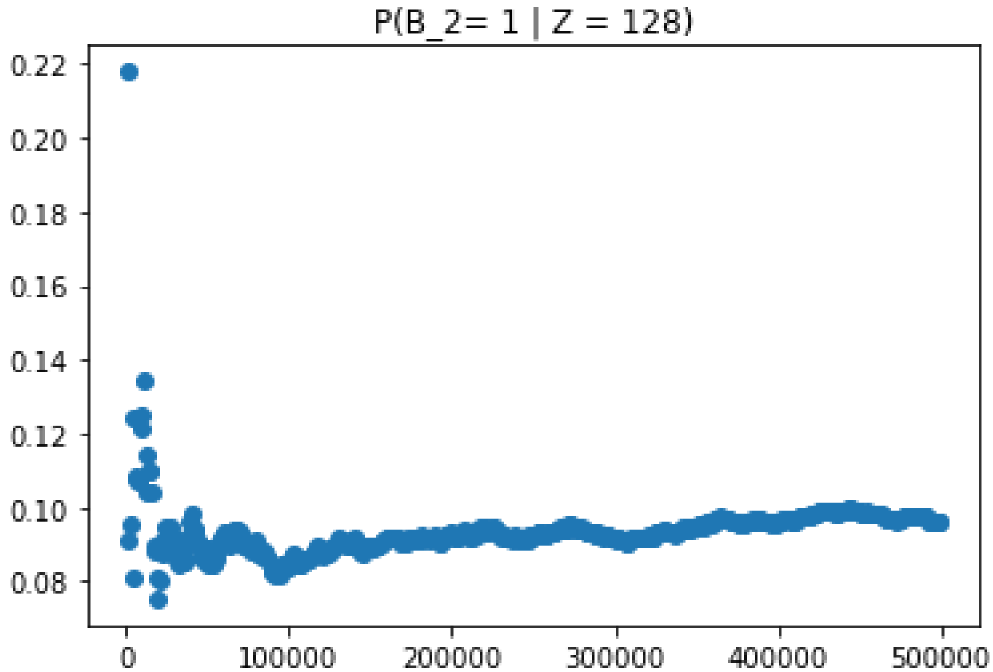
October 21, 2021

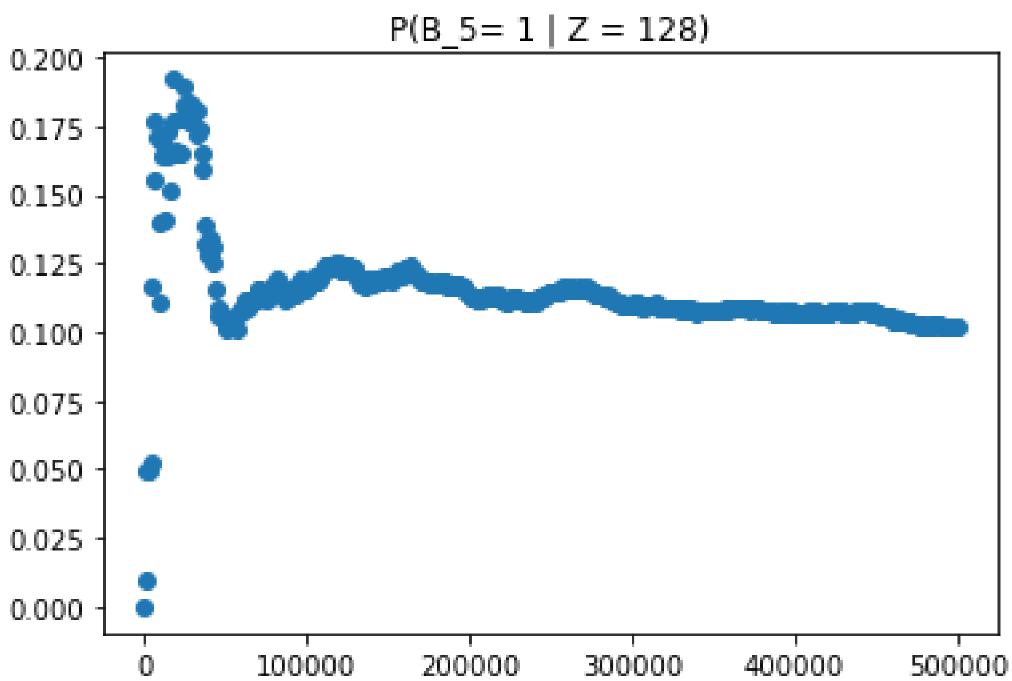
[1]:

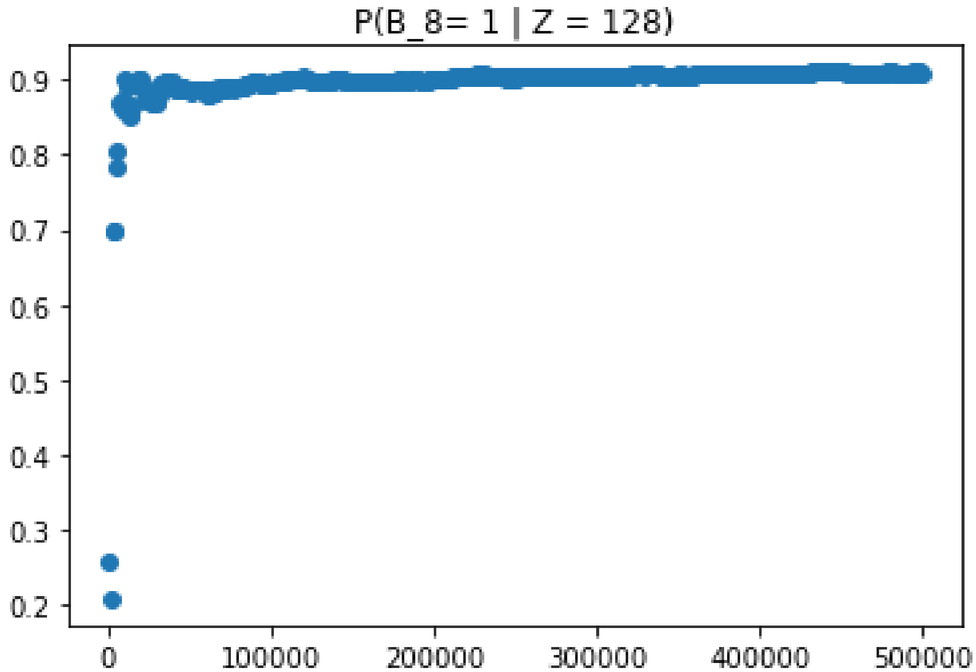
```
# Jiping Lin A15058075
# CSE 250A HW3 3.6.b 3.6.c
import random
import math
import numpy as np
import matplotlib.pyplot as plt


# 3.6.b
# [B_10, B_9, ..., B_1]
def bi_ran() -> list:
    res = []
    for i in range(10):
        res.append(random.randint(0, 1))
    return res


def fb(B: list) -> int:
    res = 0
    for i in range(len(B)):
        res += math.pow(2, len(B) - i - 1) * B[i]
    return int(res)


def pzb(z: int, B: list, alpha: float) -> float:
    return (1 - alpha) / (1 + alpha) * math.pow(alpha, math.fabs(z - fb(B)))


def trail(n, i, z, alpha):
    denom = 0
    num = 0
    for k in range(n):
        B = bi_ran()
        denom += pzb(z, B, alpha)
        num += pzb(z, B, alpha) * B[len(B) - i]
    return num / denom
```

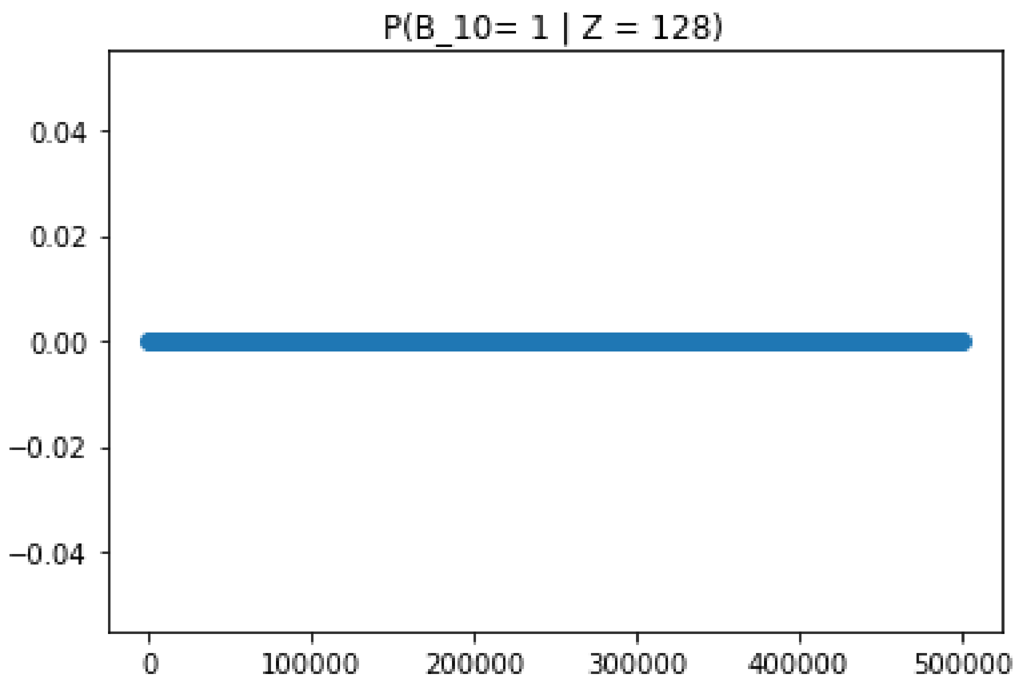
```python
def plot(k):
    x = []
    y = []
    denom = 0
    num = 0
    for i in range(500000):
        B = bi_ran()
        denom += pzb(128, B, 0.1)
        num += pzb(128, B, 0.1) * B[len(B) - k]
        if i % 1000 == 0:
            if denom == 0:
                continue
            x.append(i)
            y.append(num / denom)
    plt.scatter(x, y)
    plt.title(f"P(B_{k}" "= 1 | Z = 128)")
    plt.show()


def main():
    print("Total Epochs: 100000")
    print("P(B_2 = 1 | Z = 128) = " + str(trail(100000, 2, 128, 0.1)))
    print("P(B_5 = 1 | Z = 128) = " + str(trail(100000, 5, 128, 0.1)))
    print("P(B_8 = 1 | Z = 128) = " + str(trail(100000, 8, 128, 0.1)))
    print("P(B_10 = 1 | Z = 128) = " + str(trail(100000, 10, 128, 0.1)))
    plot(2)
    plot(5)
    plot(8)
    plot(10)


if __name__ == '__main__':
    main()
```

```
Total Epochs: 100000
P(B_2 = 1 | Z = 128) = 0.0934654915645672
P(B_5 = 1 | Z = 128) = 0.09148419592434007
P(B_8 = 1 | Z = 128) = 0.9059261103727173
P(B_10 = 1 | Z = 128) = 0.0
```


P(B_2= 1 | Z = 128)
0.22
0.20
0.18
0.16
0.14
0.12
0.10
0.08
0
100000
200000
300000
400000
500000


P(B_5= 1 | Z = 128)
0.200
0.175
0.150
0.125
0.100
0.075
0.050
0.025
0.000
0
100000
200000
300000
400000
500000


P(B_8= 1 | Z = 128)
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0
100000
200000
300000
400000
500000


P(B_10= 1 | Z = 128)
0.04
0.02
0.00
−0.02
−0.04
0
100000
200000
300000
400000
500000

[ ]: