PAGE No. |
DATE / /

250A H.W. 3

Q.3.1)

$$A_{ij} = P(X_{t+1} = j \mid X_t = i)$$

$$X_1 \longrightarrow X_2 \longrightarrow X_3 \longrightarrow \cdots X_{T-1} \longrightarrow X_T \longrightarrow X_{T+1}$$

$$P(X_{t+1} = j \mid X_1 = i) = [A^t]_{ij}$$

i) Given $[A]_{ij} = P(X_{t+1} = j \mid X_t = i)$

If $t = 1$,

$$P(X_2 = j \mid X_1 = i) = [A]_{ij}$$

Now, Assuming when $t = t$,

$$P(X_{t+1} = j \mid X_1 = i) = [A^t]_{ij}$$

We should proove that, at $t = t+1$,

$$P(X_{t+2} = j \mid X_1 = i) = [A^{t+1}]_{ij}$$

Scanned with CamScanner

LHS

$P(X_{t+2}=j \mid X_1=i) = [A^{t+1}]_{ij}$

$= \sum_a P(X_{t+2}=j, X_{t+1}=a \mid X_1=i)$

$\because$ Marginalizaⁿ

$= \sum_a P(X_{t+1}=a \mid X_1=i) \cdot P(X_{t+2}=j \mid X_{t+1}=a, X_1=i)$

$\because$ Prod. rule

Now $X_1=i \perp X_{t+2}$ given X_{t+1}

$\because$ of d(i).

$\therefore \sum_a P(X_{t+1}=a \mid X_1=i) \cdot P(X_{t+2}=j \mid X_{t+1}=a)$

$= \sum_a [A^t]_{ia} \times [A]_{aj}$

$\begin{bmatrix} [A^t]_{i1} & [A^t]_{i2} & [A^t]_{i3} & \cdots & [A^t]_{in} \end{bmatrix}$

$[A^t]_{ij}$

$\begin{bmatrix} [A]_{1j} \\ [A]_{2j} \\ \vdots \\ [A]_{nj} \end{bmatrix} \to [A]_{ij}$

$= [A^t]_{ij} \times [A]_{ij}$

$= [A^{t+1}]_{ij} =$ RHS

HP

Q 3.1)

b) We need to find $[A^t]_{ij}$.

Which is a single element of the Matrix A^t.

∴ Consider,

Step 1:-

$$\text{ith}\begin{bmatrix} [\qquad\qquad] \\ \\ \\ \end{bmatrix}\begin{bmatrix} \qquad\qquad \\ \\ \\ \end{bmatrix}\begin{bmatrix} \qquad \\ \\ \\ \end{bmatrix}$$

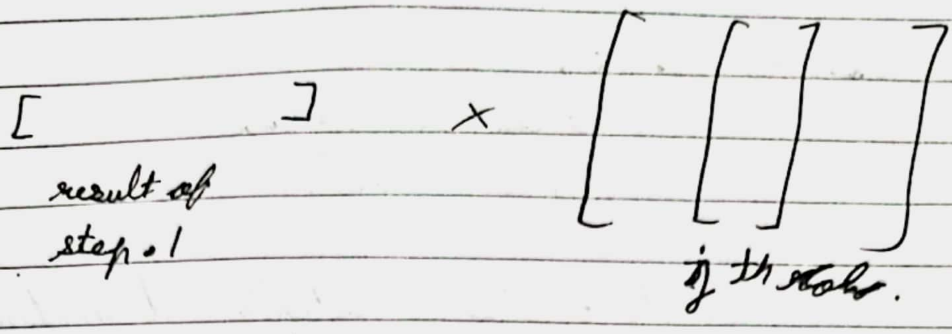
$$A \times A \times A \cdots\cdots A$$

$t - 1$ times.

We multiply the ith row $t-2$ times with matrix "A".

$$TC = n^2(t-2) = n^2 t$$

Step 2:-



Now we multiply the result of step 1 into the j th column of Mx A.

$$TC = n^2$$

- This will result in the ans of a single element that will be $[A^T]_{ij}$

$\Rightarrow$ To summarize :

$$(A_i \times A\ (t-2)\ \text{times}) \times A_j = [A^T]_{ij}$$

$$TC = n^2 t + n$$
$$= n^2 t$$
$$= m^2 t$$

Q) 3.1 c)

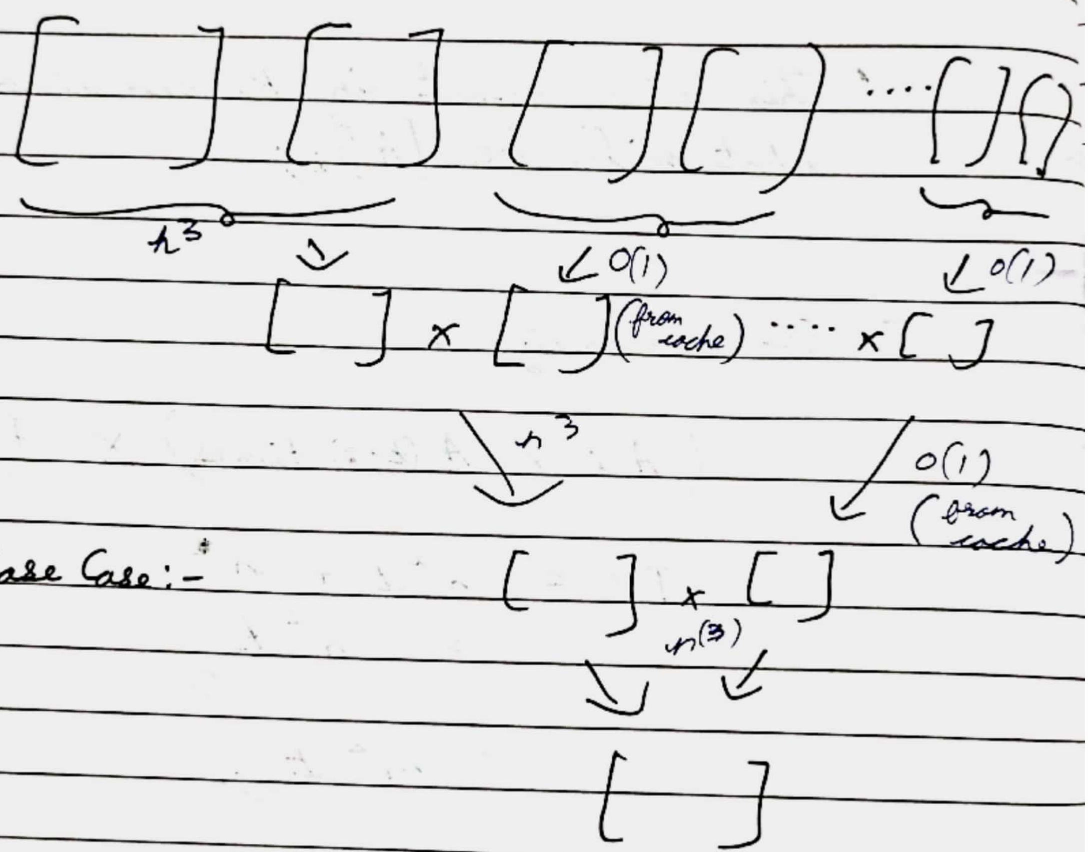
We know that,

Matrix × Matrix

$TC = n^3$.

Now we have to multiply t matrices. We can use a cache based system to drop the TC from $n^3 t$ to $n^3 \log_2 t$.

i.e.



Base Case:-

$\therefore$ we took $\log_2(t)$ times instead t.

$\therefore TC = n^3 \log_2 t$

$= m^3 \log_2 t$

PAGE No.

DATE / /

d) We kt from 3.1 b)

$$TC \text{ to find } [A^T]_{ij} = m^2 t.$$

~~In (b)'s step 1~~

In b)'s step 1 while multiplying the i^{th} row we can skip all the 0's columns while multiplying withe the matrix.

Hence $TC = t(m \times s + m \times O(1))$

skippings

$$= mst + ms$$

$$= mst$$

Following the step 2 as is,

we get the total,

$$TC = mst$$

$$= smt$$ //

Hp

PAGE No.
DATE / /

e) $P(X_1 = i \mid X_{T+1} = j)$

$$= \frac{P(X_{T+1} = j \mid X_1 = i) \quad P(X_1 = i)}{P(X_{T+1} = j)}$$

$\because$ Bayes Rule.

$$= \frac{[A^T]_{ij} \times P(X_1 = i)}{\sum_x P(X_{T+1} = j, X_1 = x)}$$

$\because$ Marginaliz^n

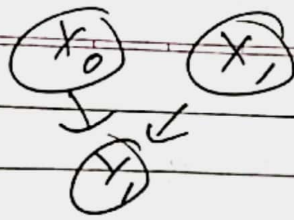
$$= \frac{[A^T]_{ij} \times P(X_1 = i)}{\sum_x P(X_{T+1} = j \mid X_1 = x) \times P(X_1 = x)}$$

$\because$ Prod-rule

$$= \frac{[A^T]_{ij} \quad P(X_1 = i)}{\sum_x [A^T]_{xj} \quad P(X_1 = x)}$$

//

Scanned with CamScanner



3.2

a) In the CPT we would have $P(Y_1 \mid X_0\, X_1)$

To find,

$$P(Y_1 \mid X_1)$$

$$= \sum_{x_0} P(Y_1, X_0 = x_0 \mid X_1) \quad \because \text{Marginaliz}^n.$$

$$= \sum_{x_0} P(X_0 = x_0 \mid X_1)\, P(Y_1 \mid X_0 = x_0, X_1) \quad \because \text{Prod. rule}$$

$$= \sum_{x_0} P(X_0 = x_0)\, P(Y_1 \mid X_0 = x_0, X_1) \quad \because \text{Marginally indep.}$$

b) To find $P(Y_1)$,

$$= \sum_{x_0} \sum_{x_1} P(X_0 = x_0, X_1 = x_1, Y_1) \quad \because \text{Marginaliz}^n$$

$$= \sum_{x_0} \sum_{x_1} P(X_0 = x_0)\, P(X_1 = x_1)\, P(Y_1 \mid X_0 = x_0, X_1 = x_1)$$

$\because$ Prod rule.
$\because$ Marginally ind.

c) To find $P(X_n \mid Y_1, Y_2, \ldots, Y_{n-1})$

$$= P(X_n) \quad \because \text{Given } E = \{\,\}, \text{ it is } \perp \text{ at } Y_n \text{ by } d\,(iii)$$

d) $P(Y_n \mid X_n, Y_1, Y_2, \dots, Y_{n-1})$

$$= \sum_{x} P(Y_n, X_{n-1} = x \mid X_n, Y_1, Y_2 \dots Y_{n-1})$$

$\because$ Marginaliz^

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

$\because$ Prod. rule

$$= \sum_{x} P(X_{n-1} = x \mid Y_1 \dots Y_{n-1}) \times P(Y_n \mid X_{n-1} = x, X_n)$$

$\because$ Given $\{\dots\}$, $X_{n-1} = x \perp\!\!\!\perp X_n$ By d(iii)

$\because$ Given X_{n-1}, $Y_n \perp\!\!\!\perp Y_1 \dots Y_{n-1}$ By d(ii)

Q.3.2) e)

$$P(Y_n \mid Y_1, Y_2 \cdots Y_{n-1})$$

$$= \sum_{x} \sum_{x_1} P(Y_n, X_{n-1} = x, X_n = x_1 \mid Y_1 \cdots Y_{n-1})$$

$\because$ marginaliz^n

$$\cancel{= \sum_{x} \sum_{x_1} P(X_n = x_1 \mid Y_1 \cdots Y_{n-1}) P(Y_n \mid X_n = x_1, Y_1 \cdots Y_{n-1}) \times P(X_{n-1} = x \mid X_n = x_1, Y_n, Y_1 \cdots Y_{n-1})}$$

$$\cancel{= \sum_{x} \sum_{x_1} P(X_n = x_1)}$$

By Prod. rule :-

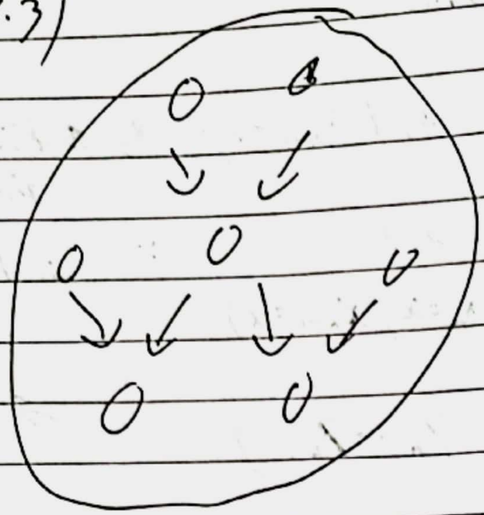
$$= \sum_{x} \sum_{x_1} P(X_n = x_1 \mid Y_1 \cdots Y_{n-1}) P(X_{n-1} = x_1 \mid X_n = x, Y_1 \cdots Y_{n-1}) \times P(Y_n \mid X_n = x, X_{n-1} = x_1, Y_1 \cdots Y_{n-1})$$

$$= \sum_{x} \sum_{x_1} \underbrace{P(X_n = x_1)}_{\because d(iii)} \underbrace{P(X_{n-1} = x_1 \mid Y_1 \cdots Y_{n-1})}_{d(iii)} \times \underbrace{P(Y_n \mid X_n = x, X_{n-1} = x_1)}_{d(ii)}$$

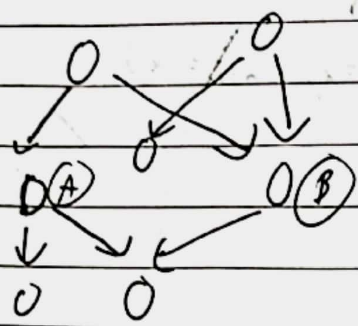
//

DATE / /

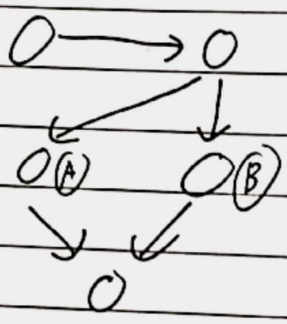
3.3)



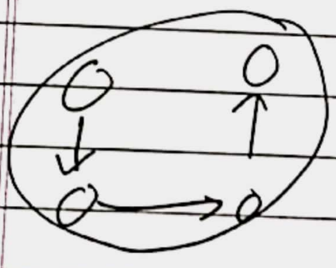
Polytree ✓



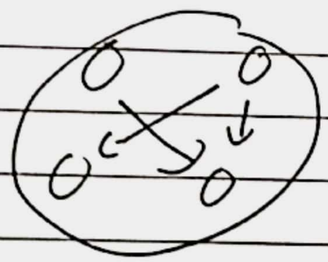
Node A & B can be clustered to make polytree.



Nod A & B can be clustered to make PT.



Polytree ✓

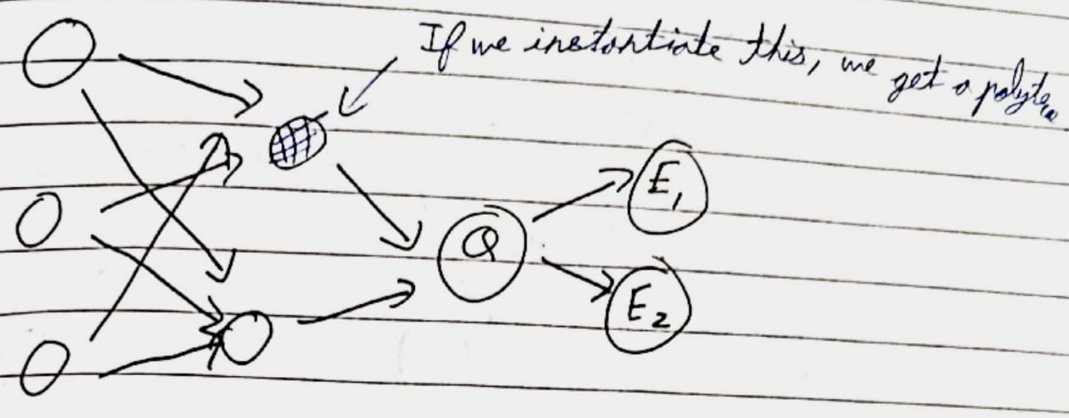


Polytree ✓

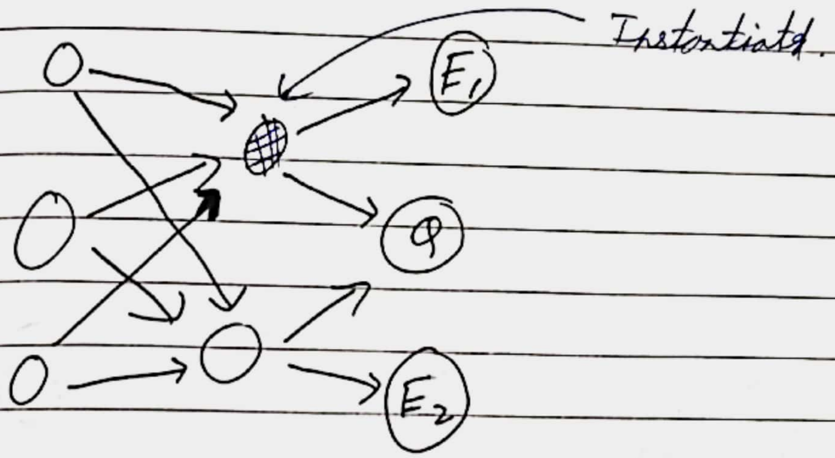
//

Q.3.4)

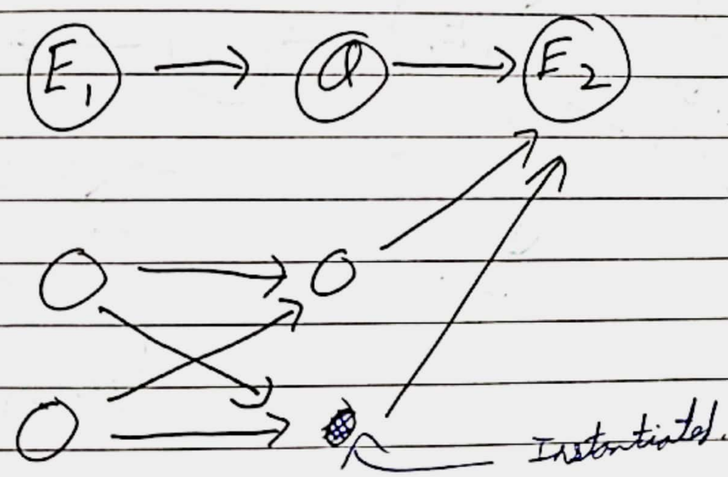
1)


If we instantiate this, we get a polytree
Q
E_1
E_2

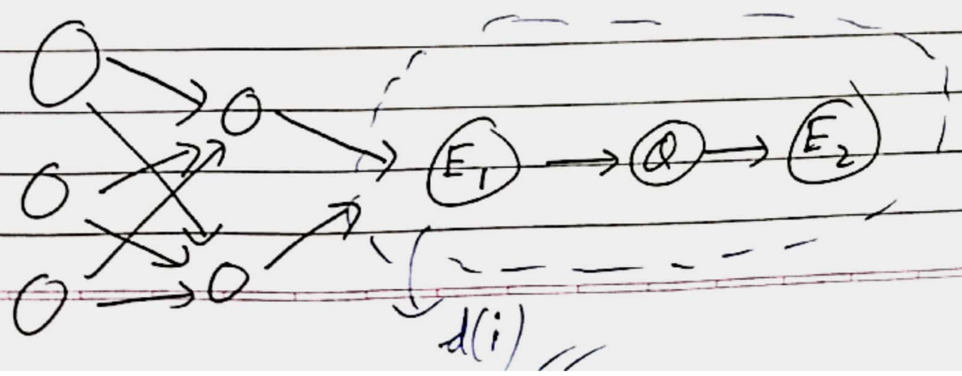
2)


Instantiated.
E_1
Q
E_2

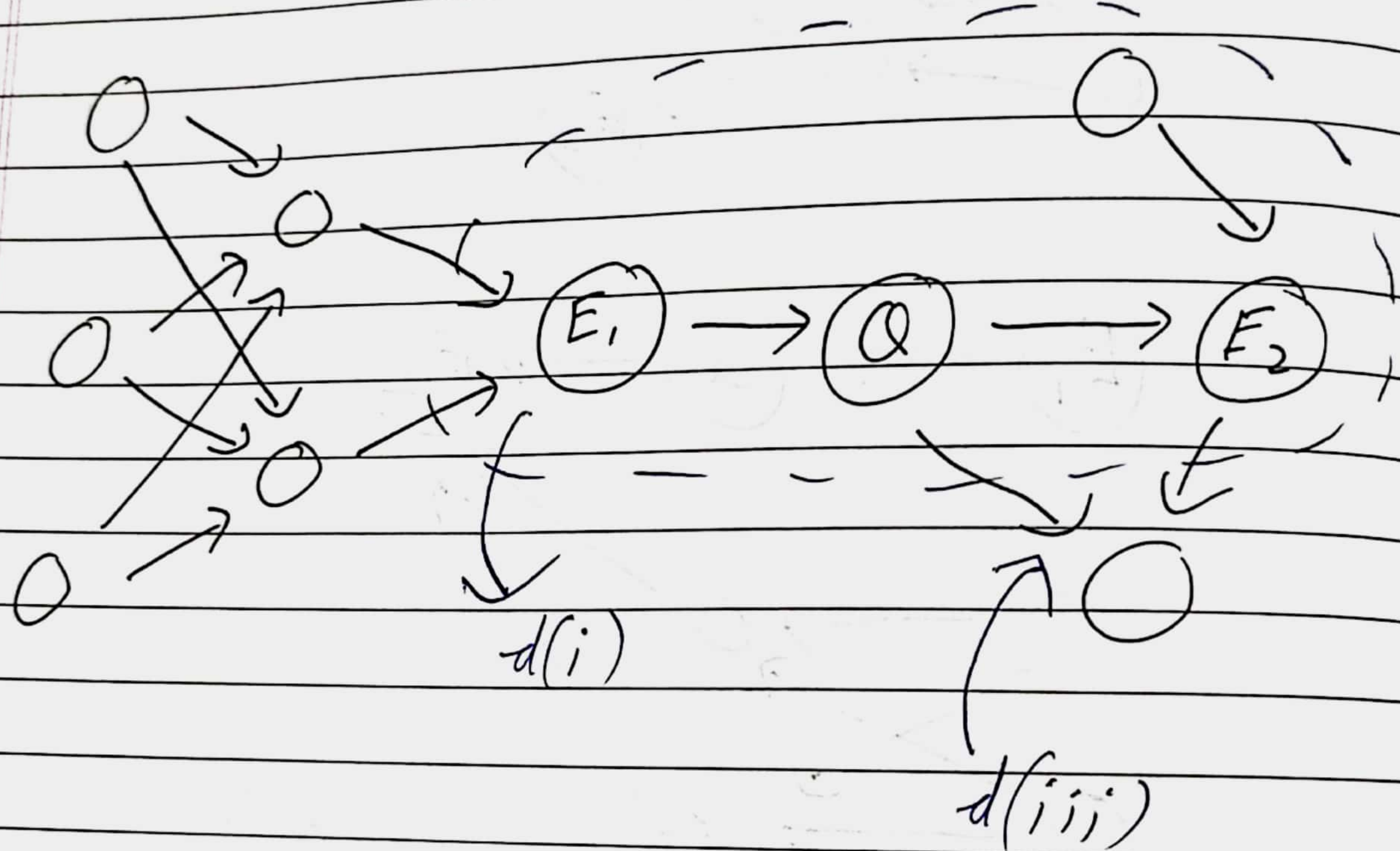
3)


E_1
Q
E_2
Instantiated.

4)


E_1
Q
E_2
d(i)

5)



Q.3.5)

$$P(z \mid B_1, B_2, \dots, B_n) = \left(\frac{1-\alpha}{1+\alpha}\right)\alpha^{|z-f(B)|}$$

$$P(B_i = 1) = \frac{1}{2}$$

$$f(B) = \sum_{i=1}^{n} 2^{i-1} B_i$$

1)

$$\sum_{z=-\infty}^{\infty} P(Z=z \mid B_1, B_2, \dots, B_n) = 1$$

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

$k = Z=z - f(B)$

When 1) $Z = -\infty$

$k \to -\infty$

2) $Z = \infty$

$k \to +\infty$

$$\therefore = \sum_{k=-\infty}^{\infty} \left(\frac{1-\alpha}{1+\alpha}\right)\alpha^{|k|}$$

$$\cancel{\sum_{k=}^{\infty}} = \frac{1-\alpha}{1+\alpha}\left(\alpha^{|k=0|} + 2\sum_{k=1}^{k=\infty} \alpha^{k}\right)$$

$$= \frac{1-\alpha}{1+\alpha}\left(1 + 2\left(\frac{1}{1-\alpha} - \alpha^{k=0}\right)\right)$$

$$\because \sum_{i=0}^{\infty} a^{i} = \frac{1}{1-a}$$

$$= \frac{1-\alpha}{1+\alpha}\left(1+\frac{2}{1-\alpha}-2\right)$$

$$= \frac{\cancel{1-\alpha}}{1+\alpha}\left(\frac{1-\cancel{\alpha}+\cancel{2}-\cancel{2}+\cancel{\alpha}\,2\alpha}{\cancel{1-\alpha}}\right)$$

$$= \frac{1+\alpha}{1-\alpha} = 1$$

//

HP

3.5)

b.) $n = 10$ $\quad i = [2, 5, 8, 10]$

$\alpha = 0.1$

For K samples,

$$P(B_i = 1 \mid Z = 128) \approx \frac{\sum_{t=1}^{K} I(b_{it}, 1)\, P(Z = 128 \mid B_1 = b_{1t} \cdots B_{10} = b_{10t})}{\sum_{t=1}^{K} P(Z = 128 \mid B_1 = b_{1t} \cdots B_{10} = b_{10t})}$$

$$= \frac{\sum_{t=1}^{K} I(b_{it}, 1) \left(\frac{1-\alpha}{1+\alpha}\right) \times \alpha^{\left|128 - \sum_{n=1}^{10} 2^{n-1} B_{nt}\right|}}{\sum_{t=1}^{K} \left(\frac{1-\alpha}{1+\alpha}\right) \times \alpha^{\left|128 - \sum_{n=1}^{10} 2^{n-1} B_{nt}\right|}}$$

$$= \frac{\sum_{t=1}^{K} I(b_i^k, 1) \times 0.1^{\left|128 - \sum_{n=1}^{10} 2^{n-1} B_{nt}\right|}}{\sum_{t=1}^{K} 0.1^{\left|128 - \sum_{i=1}^{10} 2^{n-1} B_i\right|}}$$

P.T.O.

DATE / /

∴ Ans from code :-

Bit # Prob.

"2" : 0.1 0 4 0 8 7 0 7 1 8 2 8 5 0 1 6 7

"5" : 0.0 8 9 5 5 1 9 1 3 4 8 7 5 2 8 6

"8" : 0.9 1 0 9 7 1 9 3 7 6 9 1 5 0 6 6

"10" : 0.0

//

Q) 3.6) a)

$$\begin{tikzcd} A \arrow[r] & B \arrow[r] & D \\ & C \arrow[ur] & \end{tikzcd}$$

$P(B \mid A, C, D)$

$$= \cancel{\frac{P(A,B,C,D)}{P(A)\,P(B \mid A)\,P(C \mid B,A)}}$$

$$= \frac{P(A,B,C,D)}{P(A)\;P(C \mid A)\;P(D \mid A,C)} \qquad \because \text{Prod·Rule}$$

$$= \frac{\cancel{P(A)}\;P(B \mid A)\;P(C \mid A,B)\;P(D \mid A,B,C)}{\cancel{P(A)}\;P(C \mid A)\;P(D \mid A,C)}$$

$\because P(C \mid AB)$ & $P(C \mid A) = P(C)$ By d(iii)

$$= \frac{P(B \mid A)\;\cancel{P(C)}\;P(D \mid B,C)}{\cancel{P(C)}\;P(D \mid A,C)}$$

$\because P(D \mid ABC) = P(D \mid BC)$ By d(i)

$$= \frac{P(B \mid A)\;P(D \mid B,C)}{\sum_b P(D, B=b \mid A,C)} \qquad \because \text{Marginaliz"}$$

$$= \frac{P(B \mid A)\;P(D \mid BC)}{\sum_b P(B=b \mid AC)\;P(D \mid A, B=b, C)} \qquad \because \text{Prod· Rule}$$

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

$\because P(D \mid A, B=b, C) = P(D \mid B=b, C)$ by d(i)

& $P(B=b \mid AC) = P(B=b \mid A)$ by d(iii)

b) $P(B \mid A, C, D, E, F)$

$= P(B \mid A, C, D, E)$ $\quad \because P(B \mid \overset{CDE}{\hat{A}} F) = P(B \mid \overset{CDE}{\hat{A}})$ by d(ii)

$= P(B \mid A, C, D)$ $\quad \because P(B \mid ACDE) = P(B \mid ACD)$ by d(ii)

$=$ Ans of (a)

c) $P(B, E, F \mid A, C, D)$

$= P(E, F \mid A, C, D)\; P(B \mid A, B, C, D, E, F)$ $\quad \because$ Prod rule

$\qquad$ in (b)

$= P(E, F \mid A, C) \times (b)$ $\quad \because$ by d(ii)

$= P(F \mid A, C)\; P(E \mid F, A, C) \times (b)$ $\quad \because$ Prod. rule.

$= P(F \mid A)\; P(E \mid C)\; (b)$ $\quad \because$ by d(ii)

Scanned with CamScanner

DATE / /

3.7) More likelihood weighting:-

a) Single node of evidence.

$$P(Q = q \mid E = e)$$

$$= \frac{\sum_{i=1}^{n} I(Q, q_i)\, P(E = e \mid Y = y_i, Z = z_i)}{\sum_{i=1}^{n} P(E = e \mid Y = y_i, Z = z_i)}$$

∵ from Likelihood formula.

~~b)~~ $\cancel{I(Q, q_i)}$

b) $$\frac{\displaystyle\sum_{t=1}^{n} I(Q_1, q_{1t})\, I(Q_2, q_{2t}) \times P(E_1 = e_1 \mid Q_1 = q_1, X = x_t) \times P(E_2 = e_2 \mid E_1 = e_1, Z = z_t)}{\displaystyle\sum_{t=1}^{n} P(E_1 = e_1 \mid Q_1 = q_1, X = x_t) \times P(E_2 = e_2 \mid E_1 = e_1, Z = z_t)}$$

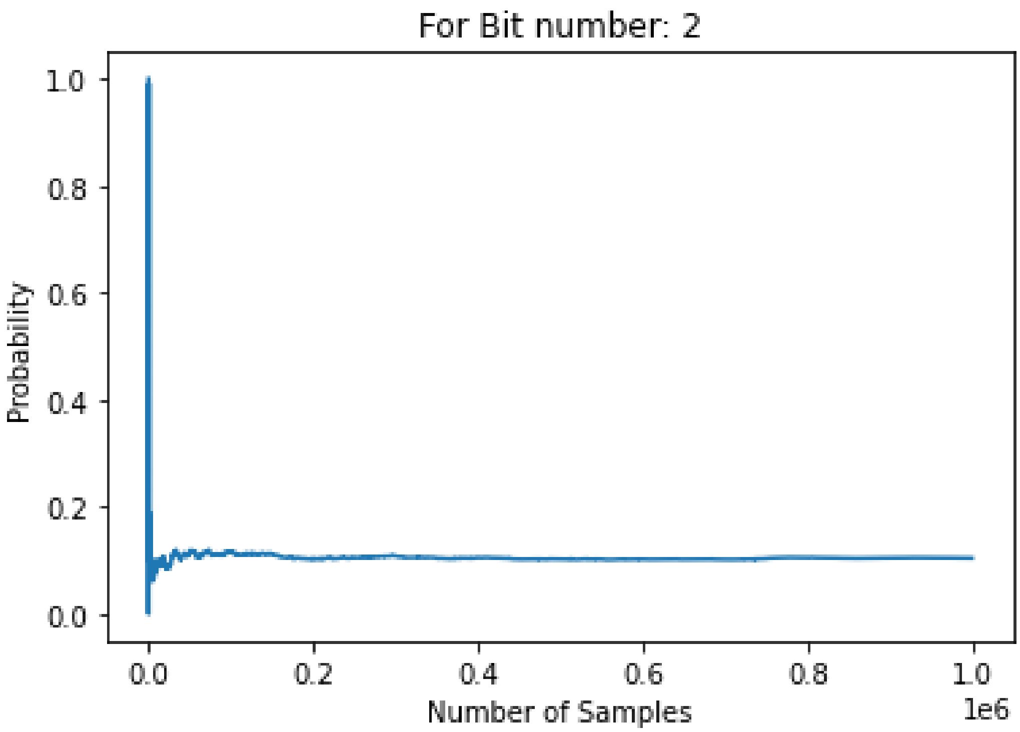
$\because$ from likelihood formula for multiple nodes.

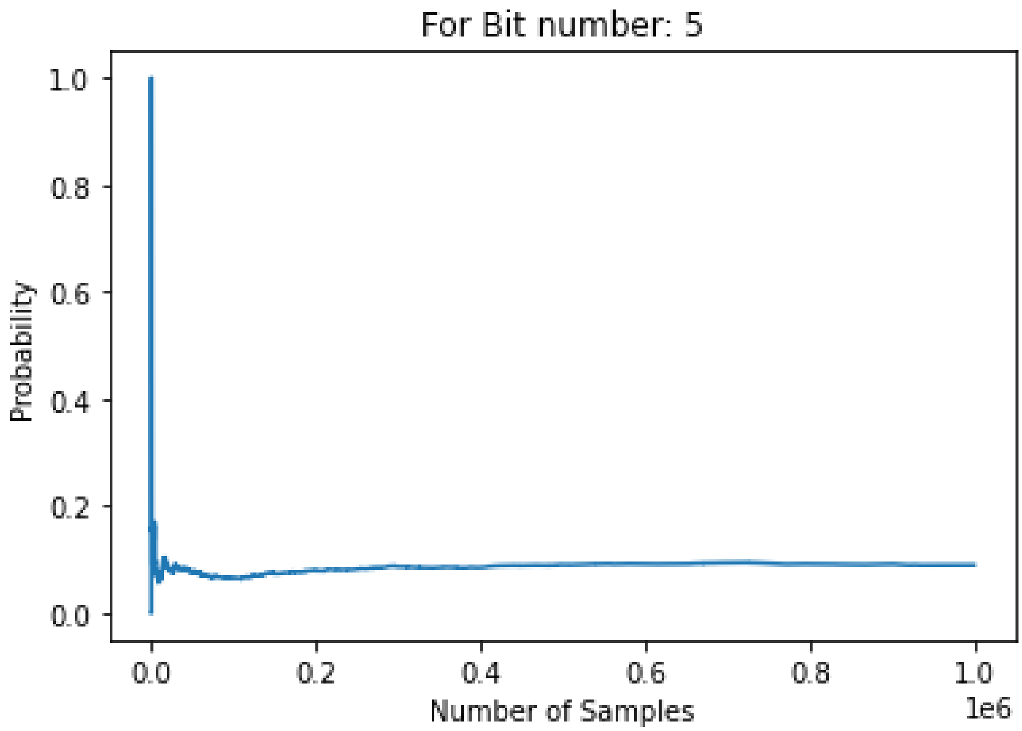
```python
from random import randint
import numpy as np
import math
from matplotlib import pyplot as plt

N = 10 # number of bits
alpha = 0.1
I = [2,5,8,10]
S = 1000000 # samples

probForEachSample = np.zeros((len(I),S))
# probForEachSample=[]

for i in range(len(I)):
  tmpNum = 0.0
  tmpDen = 0.0
  for s in range(1,S+1):
    B = []
    for b in range(1,N+1):  # random bits for this sample
      B.append(randint(0,1))
    Indicator = int(B[I[i]-1] == 1)

    tmp = 0
    for b in range(1,N+1):
      tmp+=((pow(2,(b-1)))*B[b-1])
    # print(s,tmp)
    tmp = 128 - tmp
    # print(s,tmp)
    # print("abs",abs(tmp))
    secondTerm = pow(alpha,abs(tmp))
    # print(s,"secondTerm",secondTerm)
    tmpNum += (Indicator*secondTerm)
    tmpDen += secondTerm
    # print(tmpNum,tmpDen)
    if tmpDen == 0:
      if s==1:
        probForEachSample[i,s-1] = 0
      else:
        probForEachSample[i,s-1] = probForEachSample[i,s-2]
    else:
      probForEachSample[i,s-1] = tmpNum/tmpDen
  # probForEachSample.append(tmpNum/tmpDen)

print("Finals Answers:")
for i in range(len(I)):
  print(I[i],": ",probForEachSample[i,S-1])
```

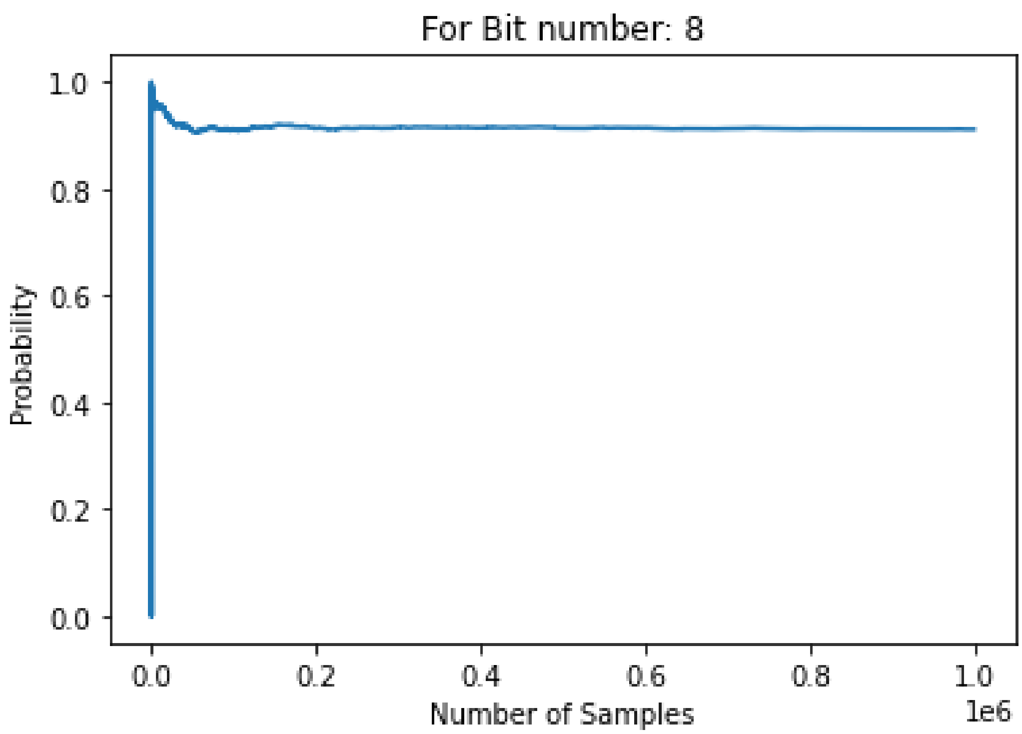
```
Finals Answers:
2 :  0.10408707182850167
5 :  0.0895519134875286
```

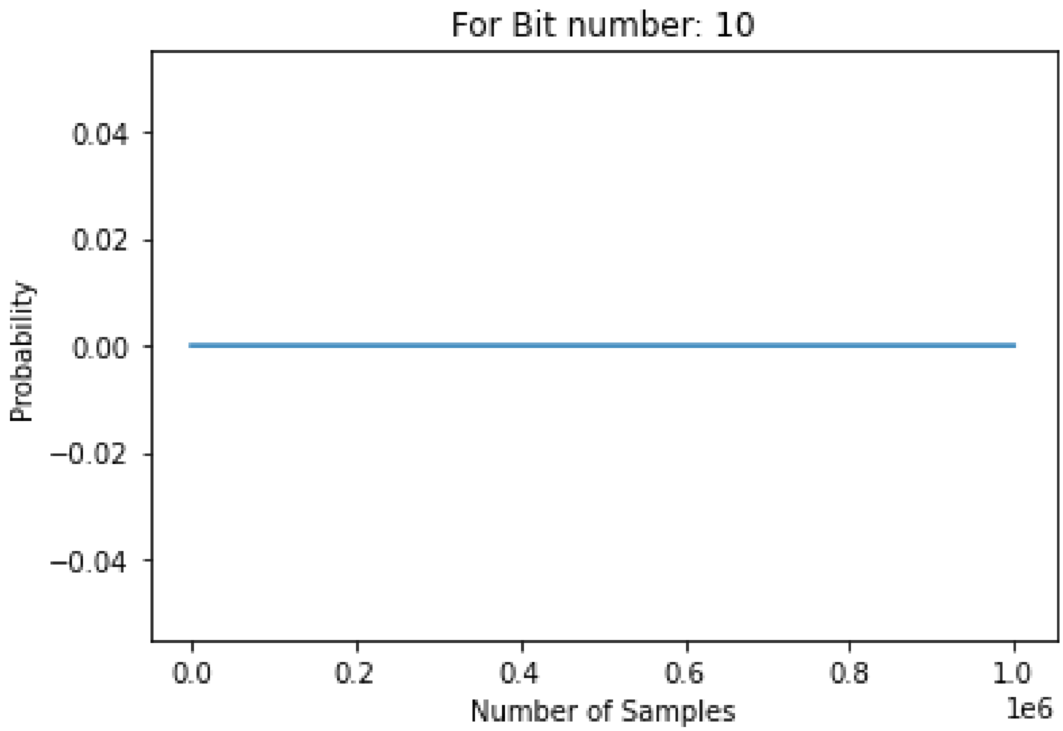
```
8 :  0.9109719376915066
10 :  0.0
```

```
# Printing graph
for i in range(len(I)):
  plt.plot(range(1,S+1),probForEachSample[i])
  plt.title("For Bit number: "+str(I[i]))
  plt.xlabel('Number of Samples')
  plt.ylabel('Probability')
  plt.show()
```




For Bit number: 5
Probability
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
1e6
Number of Samples


For Bit number: 8
Probability
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
1e6
Number of Samples


For Bit number: 10
Probability
0.04
0.02
0.00
−0.02
−0.04
0.0
0.2
0.4
0.6
0.8
1.0
1e6
Number of Samples