'1' enters 01011 000 10111 0 11 00

101 1 0011

10 101 1000 10 111 0 1100 101 1001 11

2^3 2^2 2^1 2^0

8 4 2 1

20th bit

∴ There are 11 1's in last 20 bits.

Q] How many 1's in last 15th bits of given stream after appending new bits as 111 by using PGIM algorithm.

Stream: 110111011000110101001101111 | 111

There are 10 1's in given stream in last 15 bits.

Flajolet Martin Algorithm:-

1) The Flajolet Martin Algorithm is also known as probabilistic algorithm which is mainly used to count the number of unique elements in a stream or database.

2) This algorithm was invented by Philippe Flajolet and G. Nigel Martin in 1983 and since then it has been used in various applications such as data mining and database applications.

3) The basic idea to which Flajolet-Martin algorithm is based on is to use a hash function to map the elements in the given dataset to a binary string as an estimator for the number of unique elements ~~for~~ to use as a value element.

The steps for the Flajolet-Martin algorithm are:-
1) The first step is to choose a hash function that can be used to map the elements in the database to fixed-length binary strings. The length of the binary string can be chosen based on accuracy ~~required~~ desired.
2) Next step is to apply the hash function to each data item in the dataset to get its binary string representation.
3) Next step includes determining the position of the rightmost zero in each binary string.
4) Next we compute the maximum position of the rightmost zero for all binary strings.
5) Now we estimate the no. of distinct elements in the database as 2 to the power of the maximum position of the rightmost zero which we calculated in previous step.

Q Determine distinct element using FM algorithm.
Hash function $H(x) = 6x+1$ & mod 5

Step 1:-
$h(x) = h(1) = (6x+1) \bmod 5 = 7 \bmod 5 = 2$
$h(2) = = 13 \bmod 5 = 3$
$h(3) = = 19 \bmod 5 = 4$
$h(4) = = 25 \bmod 5 = 0$

Step 2:- Binary representation (3 bit)

$h(1) = 2 = 010 = 1$

$h(2) = 3 = 011 = \boxed{0}$

$h(3) = 4 = 100 = \boxed{2}$

$h(4) = 0 = 000 = \boxed{0}$

Step 3:- Count each 0 after one (1) from binary representation of step 2.

$h(1) = 1$

$h(2) = 0$

$h(3) = 2$

$h(4) = 0$

Step 4:- Find maximum value from step 3

$h(3) = 2$

$\therefore x = 2$

Step 5:- Distinct element

$R = 2^x = 2^2 = 4$

$\boxed{R = 4}$

∴

8] Page Rank Numerical

$PR(A) = (1-d) + d\,(PR(T_i)/C(T_i) + \ldots + PR(T_n)/C(T_n))$

$PR(A)$ = Page rank A

$PR(T_i)$ = PR of pages T_i which link to page A

$C(T_i)$ = no. of outbound links on page T_i

d = damping factor

↓

Teleport factor