1 Bloom Filter (Cont.)

After inserting $m$ items into $n$ slots with $k$ hash functions

- $P[A[i] = 0] = (1 - \frac{1}{n})^{km} \approx e^{-\frac{km}{n}}$
- $P[A[i] = 1] = 1 - (1 - \frac{1}{n})^{km} \approx 1 - e^{-\frac{km}{n}}$

We want to obtain the False Positive Rate

\[
P[\text{lookup} (y) = \text{Yes}] = P[A[h_1 (y)] = 1 \cap \cdots \cap A[h_k (y)] = 1] = \prod_{i=1}^{k} P[A[h_i (y)] = 1] = \left(1 - e^{-\frac{km}{n}}\right)^k
\]

for fixed $m$ and $n$, minimize the probability, we get

\[
k = \frac{n}{m} \ln 2
\]

and corresponding approximated False Positive rate $\left(1 - e^{-\ln 2}\right)^k$.

Following table gives some related results

<table>
<thead>
<tr>
<th>$n/m$</th>
<th>$k$</th>
<th>False Positive rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>6</td>
<td>2.16%</td>
</tr>
<tr>
<td>16</td>
<td>11</td>
<td>0.04587%</td>
</tr>
<tr>
<td>32</td>
<td>22</td>
<td>$2.1 \times 10^{-5}$%</td>
</tr>
</tbody>
</table>
2 Quotient Filter

**Definition 1.** The quotient filter is a compact hash table. The hash function generates a \( p \)-bit fingerprint. The \( r \) least significant bits is called the remainder and \( q = p - r \) most significant bits is called quotient. And the hash table has \( 2^q \) slots.

Three additional bits are used to reconstruct a slot’s fingerprint:

- occupied: set when a slot is the canonical slot for some key stored
- continuation: set when a slot is occupied but not by the first remainder in a run
- shift: set when the remainder in the slot is not in its canonical slot

The following graph shows how to insert in quotient filter.