Bloom Filters

Learning Objectives

1. Know the optimal hash formula for Bloom Filters

Optimal Hash Formulation

$$\left(1-\left(1-\frac{1}{m}\right)^{nk}\right)^{k}$$

Derivative is zero when $k^* = \ln 2 \cdot \frac{m}{n}$

Optimal Parameter

$$k^* = \ln 2 \cdot \frac{m}{n}$$

Given any two values, we can optimize the third

n = 1000 items	m = 1000 bits	k =
m = 1000 bits	k = 5 hashes	n =
k = 5 hashes	n = 1000 items	m =

Optimal Parameter

$$k^* = \ln 2 \cdot \frac{m}{n}$$

Given any two values, we can optimize the third

n = 1000 items	m = 1000 bits	k ≈ .69
m = 1000 bits	k = 5 hashes	n ≈ 138.6
k = 5 hashes	n = 1000 items	m ≈ 7213.5