=: Trade-off b/w distance & rate :=

Motivating Question 1: What is the largest rate R that a code with distance d can have?

Example: $C_{3,rep}$: $[12, 4, 3]_2 \Rightarrow R = \frac{1}{3}$ & $d = 3$.

$C_H : [7, 4, 3]_2 \Rightarrow R = \frac{4}{7}$ & $d = 3$.

$\Rightarrow C_H$ is a better code compare to $C_{3,rep}$.

Q1: Does we have a binary code of block length 7 & distance 3 but rate $> 4/7$?

Hamming Bound: Our first trade-off b/w redundancy (in the form of dimension of a code) & its error-correction capability (in the form of its distance):

Def^n (Hamming ball): $\forall x \in \Sigma^n$ with $|\Sigma| = q$ & $e \in \mathbb{N}$

$$B(x,e) = \{ y \in \Sigma^n \mid \Delta(x,y) \le e \}.$$

Observation 1:

$$|B(x,e)| = \sum_{i=0}^{e} \binom{n}{i} (q-1)^i$$

Observation 2: $\forall x, y \in \Sigma^n$, & $e \in \mathbb{N}$

$$|B(x,e)| = |B(y,e)|$$

$\Rightarrow |B(x,e)| = |B(0,e)| = \#$ n-length strings/words of weight $\le e$.

We use $Vol_q(e,n)$ to denote $|B(x,e)|$.

[Called Volume of Hamming Ball].

- $Vol_2(e,n) = \sum_{i=0}^{e} \binom{n}{i}$

- $Vol_2(1,n) = \binom{n}{0} + \binom{n}{1} = n+1.$

We prove an upper bound on the dimension - every binary code of distance 3.

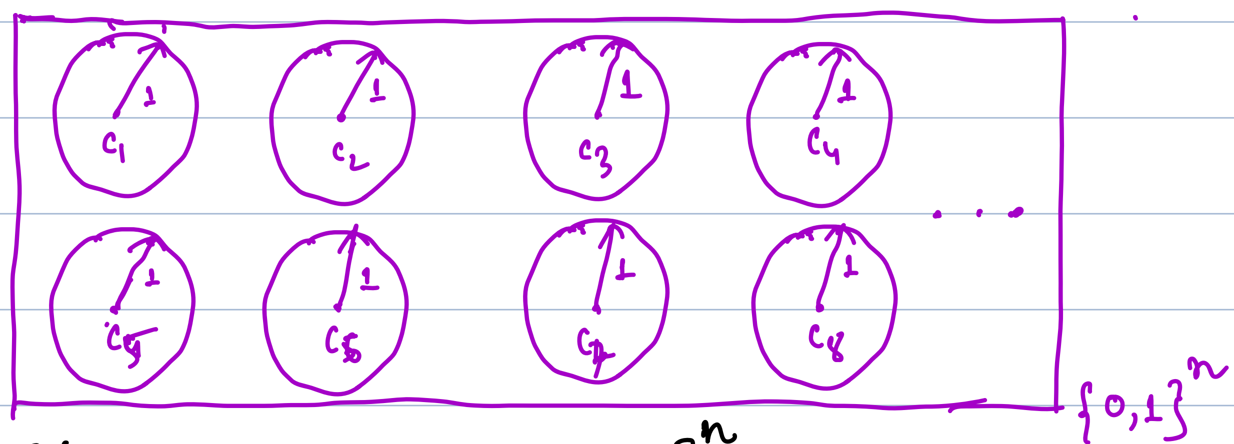
Theorem 1: For every $(n,k,3)_2$ code,

$$k \lesssim n - \log_2(n+1) \qquad \text{[negative result]}$$
$$= n - \log_2 Vol_2(1,n)$$

Theorem 1 $\Rightarrow$ for $n=7$, C_H is the best possible code of block length 7 & dist: 3.

Proof of Thm 1: Let C be an $(n,k,1)_2$ code. –

$\Rightarrow \forall c_1 \neq c_2 \in C, \quad B(c_1,1) \cap B(c_2,1) = \phi \quad \cdots \text{(i)}$



$$|B(x,1)| = n+1 \quad \forall x \in \{0,1\}^n$$

$$\bigcup_{c \in C} B(c,1) \subseteq \{0,1\}^n \Rightarrow \left| \bigcup_{c \in C} B(c,1) \right| \leq 2^n$$

$$\Rightarrow \sum_{c\in C} |B(c,1)| \leq 2^n \text{ ---- (from the equation (1))}$$

$$\Rightarrow |C|\cdot(n+1) \leq 2^n \Rightarrow k = \log|C| \leq n - \log_2(n+1).$$

Extending Binary symbols to arbitrary alphabets. & from distance 3 to arbitrary distance.

Generalized Hamming Bound:

Thm 2: $\forall\ (n,k,d)_q$ code C,

$$k \leq n - \log_q\left(\sum_{i=0}^{\lfloor\frac{d-1}{2}\rfloor}\binom{n}{i}\cdot(q-1)^i\right).$$ [negative result]

$$= n - \log_q Vol\left(\left\lfloor\frac{d-1}{2}\right\rfloor, n\right).$$

Proof of Thm 2: let $e = \lfloor\frac{d-1}{2}\rfloor. \Rightarrow$

$\forall c\in C,\quad |B(c,e)| = \sum_{i=0}^{e}\binom{n}{i}\cdot(q-1)^i$ &

$\forall c_1, c_2 \in C,\quad B(c_1,e)\cap B(c_2,e) = \Phi.$

$$\bigcup_{c\in C} B(c,e) \subseteq \Sigma^n \Rightarrow \left|\bigcup_{c\in C} B(c,e)\right| \leq q^n$$

$$\Rightarrow \sum_{c\in C} |B(c,e)| \leq q^n$$

$$\Rightarrow q^k\cdot \sum_{i=0}^{e}\binom{n}{i}\cdot(q-1)^i \leq q^n$$

$$\Rightarrow \quad k \le n - \log_q \sum_{i=0}^{e} \binom{n}{i} (q-1)^i$$

$\Rightarrow \forall\ (n,k,d)_q$ code C,

$$R \lesssim 1 - \frac{\log_q \left(\sum_{i=0}^{e} \binom{n}{i} (q-1)^i \right)}{n}.$$

$$= 1 - \frac{Vol_q(\lfloor \frac{d-1}{2} \rfloor, n)}{n}.$$

Def$\overset{n}{\text{=}}$ (Perfect Code): Any code that achieves Hamming bound is called the perfect code.

- Optimal trade-off b/w redundancy & error-correction capability for code-families:

Code family: $C = \{C_i\}_{i \in \mathbb{N}}$ where $C_i = (n_i, k_i, d_i)_q$

& relative distance $\delta(C) = \lim_{i \to \infty} \frac{d_i}{n_i}$, when exists

rate of C, $R(C) = \lim_{i \to \infty} \frac{k_i}{n_i}$, when exists.

Motivating Question 2: What is the optimal trade-off b/w $R(C)$ & $\delta(C)$ that can be achieved by some code family C?

We try to answer this question. To do that first we need introduce a fundamental function, called entropy function, which plays a crucial role in studying trade-off b/w $R(C)$ & $\delta(C)$.

Defn (q-ary Entropy Function): Let q be a positive integer ≥ 2. Then $\forall x \in (0,1)$,

$$H_q(x) = x \cdot \log_q(q-1) - x \cdot \log_q x - (1-x)\log_q(1-x).$$

Example: 1. $q=2$, $H_2(x) = -x\log_2 x - (1-x)\log_q(1-x)$.

Shanon's Entropy Function: the entropy of the distribution over $\{0,1\}$ that selects 1 w.p. x & 0 w.p. $(1-x)$.

also used to quantify "randomness" in a distribution.

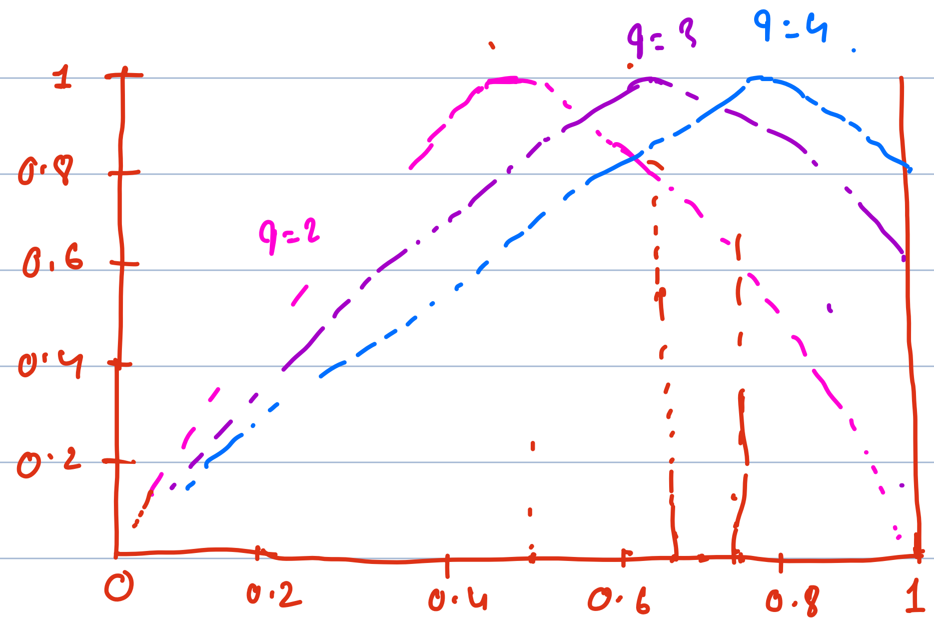
What is the max^m value of $H_2(x)$? At what point achieves?
1. ½.

For $H_2(x)$, we drop subscript & denote it by $H(x)$.

2. $q=3$: What is the max^m value of $H_3(x)$ & at what point?
1 ⅔.

In general, what is the maximum value of $H_q(x)$ & at what point?
1 $1-1/q$.

[How? Exercise!]



For $q > 2$ and $x \in (0,1)$ consider the following distribution defined on the set of q elements $\{a_1, a_2, \ldots\ldots, a_q\}$ as follows: pick a_1 with probability $(1-x)$ &

pick a_i for $i \in \{2, 3, \ldots, q\}$ w.p. $x/q-1$.

Then, Shanon Entropy of this distribution w.r.t q-base logarithm:

$$= -(1-x)\log_q(1-x) - (q-1)\cdot\frac{x}{(q-1)}\cdot\log_q\frac{x}{q-1}$$

$$= -(1-x)\log_q(1-x) - x\log_q x + x\log_q(q-1)$$

which is the q-ary entropy function $H_q(x)$ defined above!

Since Shanon Entropy achieves maximum when we define a uniform distribution on $\{a_1, a_2, \ldots, a_q\}$.

Question is can we define a uniform distribution on $\{a_1, a_2, \ldots, a_q\}$ using the above proceduaes.

Answer: Yes! assume $x = 1 - \frac{1}{q}$. Then observe that the above distribution produces a unifrom distribution on $\{a_1, a_2, \ldots, a_q\}$.

Therefore, max^m value of $H_q(\cdot)$ is 1 & it is achieved at $x = 1 - \frac{1}{q}$.

[Thanks to Aranya for giving such a nice meaning to the q-ary Entropy function].