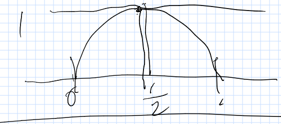
3/9/21

entropy (Shanon theorem)

$H(p) = -p \lg p - (1-p) \lg(1-p)$

$\log_2$ binary entropy function



$X \in \{1,2,\dots,n\}$

$H(X) = -\sum_{i=1}^{n} p_i \lg p_i$

$X \in \{1,\dots,n\}$

X is uniform

$H(X) = -\sum_{i=1}^{n} \text{Prob}[X=i] \lg \text{Prob}[X=i]$

$= -\sum \frac{1}{n} \lg \frac{1}{n} = \sum_{i=1}^{n} \frac{1}{n} \lg n$

$= \lg n.$

$X \in \{1,2,\dots,n\}$

distribution of X that maximize the entropy?

$H(X=1 \text{ with probability } 1) = 0$

$= 1 \cdot \lg 1 - \sum_{i=1} 0 \cdot \lg 0$

Lemma

X and Y two random variables that are independent then

$Z = (X, Y)$

$H(Z) = H(X \quad H(Y)$

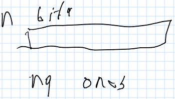
Lemma

n: integer
$q \in [0,1]$
nq is an integer

$$\frac{2^{nH(q)}}{\boxed{n+1}} \le \binom{n}{nq} \le 2^{nH(q)}$$

$$H(q) = q \lg \frac{1}{q} + (1-q) \lg \frac{1}{1-q}$$

$$\binom{n}{nq} =$$



$$q$$

$$\boxed{nH(q)}$$

---

proof

$$\binom{n}{nq} q^{nq} (1-q)^{n-nq} \le (q+(1-q))^n = 1$$

$$(x+y)^m = \sum_{i=0}^{m} \binom{m}{i} x^i y^{m-i}$$

$$\binom{n}{nq} q^{nq} (1-q)^{n(1-q)} \le 1$$

$$\binom{n}{nq} \le \frac{1}{q^{nq} (1-q)^{(1-q)n}}$$

$$= 2^{\left(\lg \frac{1}{q}\right) nq + \left(\lg \frac{1}{1-q}\right) n(1-q)}$$

$$= 2^{nH(q)}.$$

■

Extracting randomness

$X \in \{0, \dots, 2^m-1\}$
uniform

X write in base 2 and output the bits

$X \in \{0, \dots, n-1\}$ n is not a power of 2.

$x_1, x_2, x_3, \dots \sim$ distribution(X)

$ext(x_1)$ $ext(x_2)$ ...

0, 1, 2

0	0
1	1
2	ε

$\frac{1}{3}$

$ext(X_1, X_2) =$

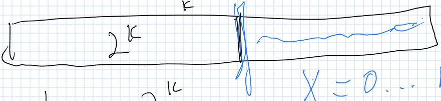
	0	1	2
0	000	001	010
1	011	100	101
2	110	111	~~X~~ ε

$\frac{1}{9}$

$X \in \{0, \dots, n-1\}$ uniform

extract(X)

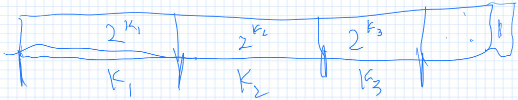
$\max_k 2^k \le n$



$X = 0 \dots n-2^k$

$1 \dots 2^k$

$0 \dots 2^k-1$



Thm

X is chosen uniformly from $\{0, \ldots, m-1\}$, then the extraction function outputs at least

$$\lfloor \log m \rfloor - 1 = \lfloor H(X) \rfloor - 1$$

$\boxed{H(X) \text{ upper bound}}$

---

q: probability for head
flip it n times

$$\boxed{nH(q)}$$

$S = s_1 s_2 \ldots s_n$ = input binary string

Y = # of 1s in S

S_k = all binary strings with k ones of len n.

$$|S_n| = \binom{n}{k}$$

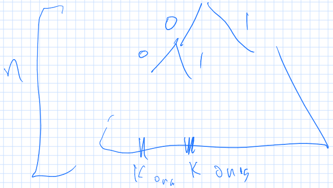
$T \in S_n$

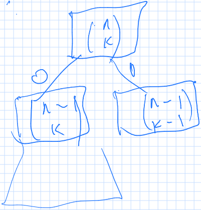
$$P[s_1 s_2 \ldots s_n = T \mid s_1 s_2 \ldots s_n \in S_k] = \frac{1}{|S_k|}$$

$$= \frac{1}{\binom{n}{k}}.$$

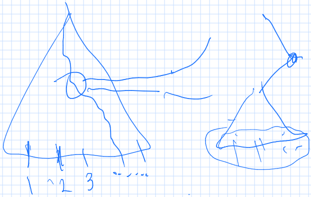
$s_1, \ldots, s_n$

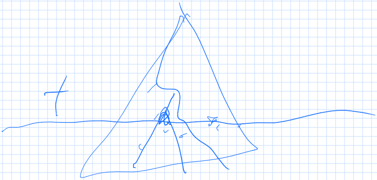
$$S_k = \left\{ \begin{array}{l|l} 00 \cdots 0\,11111 & 1 \\ 0 \cdots & 2 \\ \vdots & \vdots \\ & C^n_k \end{array} \right\}$$





$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$





$$\binom{n}{k} \leq 2^n$$ n bits

$O(n^2)$

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$$

$O(n^2 \log^2 n)$

$n^{O(1)}$

Thm

The above extraction procedure extracts $\geq n(1-\delta)H(q)$ $\delta \in (0,1$ unbiased random bits.

And one can not extract more than $nH(q)$ bits (unbiased).

$Ext(x)$ = binary string output on the input x

$$P\left[Ext(x)=y \,\middle|\, |y|=t\right] = \frac{1}{2^t}$$

$X \Rightarrow y = Ext(x)$

$$\underbrace{2^{|Ext(x)|} P[X=x]}_{\leq 1} \leq 2^{|Ext(x)|} P\left[Y=Ext(x)\right]$$

$$\leq 2^{|Ext(x)|} \frac{1}{2^{|Ext(x)|}}$$

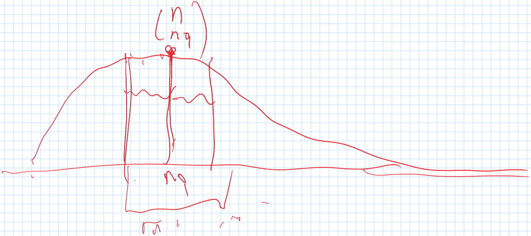
$$= 1$$

$$\boxed{|Ext(x)| \leq \lg \frac{1}{P[X=x]}}$$

$$E\left[\#\text{ rand bits extracted}\right] = \sum_x P[X=x]\,|Ext(x)|$$

$$\leq \sum_x P[X=x] \lg \frac{1}{P[X=x]}$$

$\leq H(X) \simeq nH(q),$



$n(H(q) - r)$