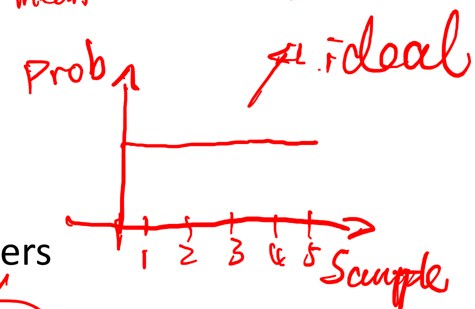
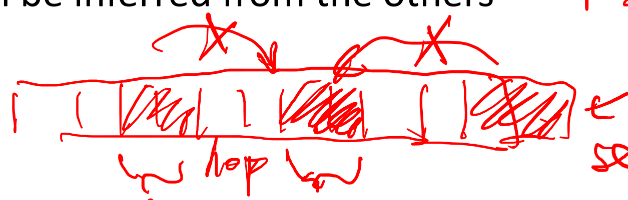
Properties of Random Numbers

- Randomness
  - Uniformity
    - distribution of bits in the sequence should be uniform

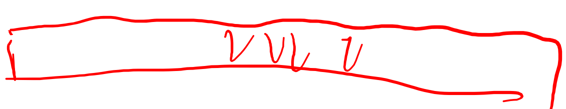
why



  - Independence
    - no one subsequence in the sequence can be inferred from the others
- Unpredictable
  - satisfies the "next-bit test“



any



1 0 1 0 0 1 0 1

consecutive

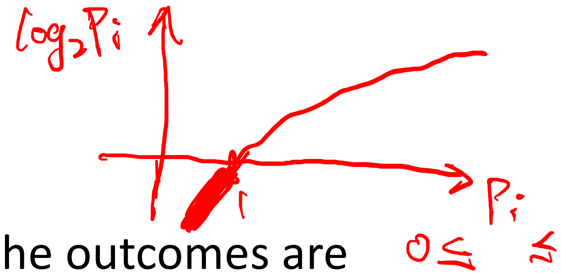
Markov Process

sub sequence

$P(AB) = P(A) \cdot P(B)$

independent

Entropy



- A measure of uncertainty
  - In other words, a measure of how unpredictable the outcomes are
  - High entropy = unpredictable outcomes = desirable in cryptography
  - The uniform distribution has the highest entropy (every outcome equally likely, e.g. fair coin toss)
  - Usually measured in bits (so 3 bits of entropy = uniform, random distribution over 8 values)

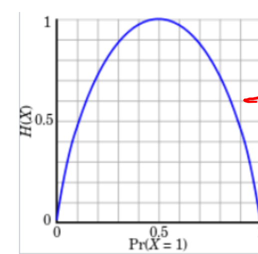
$n =$

≥ 0 $\quad < 0$ probability of value i

entropy

$$H = -\sum_i p_i \boxed{\log_2(p_i)}$$

Entropy of an information source



$n = 2.$

$$H = -\sum_{i=1}^{n=8} P_i \log_2(P_i) \checkmark$$

$$= -\left[\frac{1}{8}\cdot\log_2\frac{1}{8} + 0\cdot + \frac{1}{16}\log_2\frac{1}{16} + \cdots \frac{3}{16}\log_2\frac{3}{16}\right]$$

1 value — 2 values (0) — 3 value — 8 value

$$= -\left[-\frac{3}{8} - \frac{4}{16} - \frac{2}{4} - \frac{3}{8} - 0.45 - \frac{4}{16} - 0.45\right]$$

$$= 2.234$$

Ideal
$$H = -\sum_{i=1}^{8}\frac{1}{8}\log_2\frac{1}{8}$$
$$= -8\cdot\frac{1}{8}\log_2\frac{1}{8}$$
$$= -8\cdot\frac{1}{8}(-3) = 3$$

$$\frac{dH}{dP_i} = 0 \Rightarrow P_i^* = \frac{1}{n}$$

$$\sum_{i=1}^{n} P_i = 1$$

t period

C_i

$8 = 2^3$

$H_2 = 1.24$

prob. $P_i = \frac{C_i}{\sum_{i=1}^{8} C_i}$, $0 \le \cdot < 1$

$\sum_i P_i = 1$

empirical — $t \to \infty$ — true — $t \to \infty$

data source random

1 2 3
0 5 6
7 8

value		Prob
1	C_1	1/8
2	C_2	0
3	C_3	1/16
4		1/4
5		1/8
6		3/16
7		1/16
8		3/16

$\partial \to$ partial derivative

$$\frac{dH}{\partial p_i} = -\sum_{i=1}^{8}\left[\log_2(p_i) + p_i \cdot \frac{d\log_2 p_i}{dp_i}\right] \qquad \frac{d\log_a x}{dx} = \frac{1}{x} \rightarrow \text{search}$$

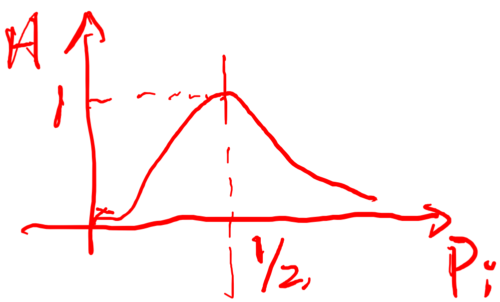
$$= 0$$

$$p_i = \frac{1}{n}$$

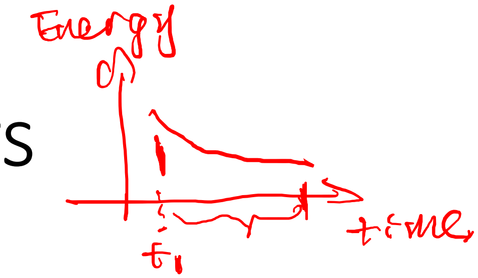
$$-\sum_{i=1}^{2} \frac{1}{2}\cdot\log_2\frac{1}{2}$$

$$2^1 = 2$$

$n=2$

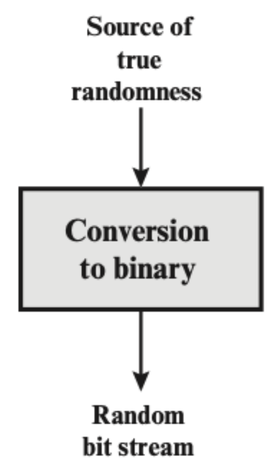


True random numbers generators



- Several sources of randomness – natural sources of randomness
  - decay times of radioactive materials
  - electrical noise from a resistor or semiconductor → thermal noise
  - radio channel or audible noise
  - keyboard timings
  - disk electrical activity
  - mouse movements
  - Physical unclonable function (PUF) → Prof. Zhang
- Some are better than others

electron movement



(a) TRNG

Combining sources of randomness

- Suppose r1, r2, …, rk are random numbers from different sources. E.g.,

  r1 = electrical noise from a resistor or semiconductor

  r2 = sample of hip-hop music on radio

  r3 = clock on computer

  b = $\boxed{r1 \oplus r2 \oplus \ldots \oplus rk}$

  If any one of r1, r2, …, rk is truly random, then so is b

  output b

  Many poor sources + 1 good source = good entropy

Pseudorandom Number Generators (PRNGs)

- True randomness is expensive
- **Pseudorandom number generator** (**PRNGs**): An algorithm that uses a little bit of true randomness to generate a lot of random-looking output
  - Also called **deterministic random bit generators** (**DRBGs**)
- PRNGs are deterministic: Output is generated according to a set algorithm
  - However, for an attacker who can't see the internal state, the output is *computationally indistinguishable* from true randomness