TTIC 31200/CMSC 37220

Information & Coding Theory

- TA: Kavya Ravichandran
- Discussion: F 2-3 pm
- Office hours: Th 12:30 - 1:30

- See course page for notes/resources
- 4 Homeworks (60%) + Take-home final (40%)

- Pre-reqs: probability, basic analysis, linear algebra

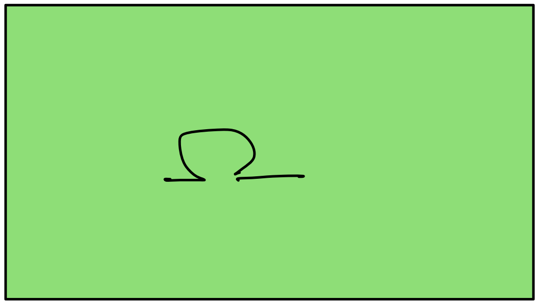
Contents

Information Theory $\simeq 75\%$ (mathematical perspective)

Coding Theory $\simeq 25\%$ (error correction)

- basic concepts + applications
- statistics
- error correcting codes
- TCS applications / miscellaneous topics / quantum information

Recap: Random Variables & Distributions



(discrete) probability space Ω

measure $\mu : \{\text{subsets of } \Omega\} \to [0,1]$

Random variable

$X : \Omega \to \underbrace{\mathbb{R}}$

set of "values" for X

Distribution

$P(X)$

"Event" Subset of Ω

eg. $\mathbb{P}[X = a] = 0.3$

$\mathbb{P}[X = b] = 0.7$

Notation

X, Y, Z .... random variables

x, y, z ..... values

$\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ .... sets of possible values (support)

P, Q ... distributions

p, q ... actual probabilities

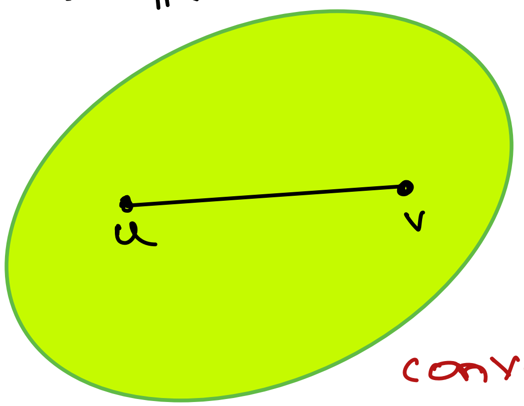
$\mathbb{P}[\text{event}]$ ...

→ when well-defined.

$$\mathbb{E}[X] = \sum_{a \in \mathcal{X}} a \cdot \mathbb{P}[X = a]$$

Convexity

$S \subseteq \mathbb{R}^d$

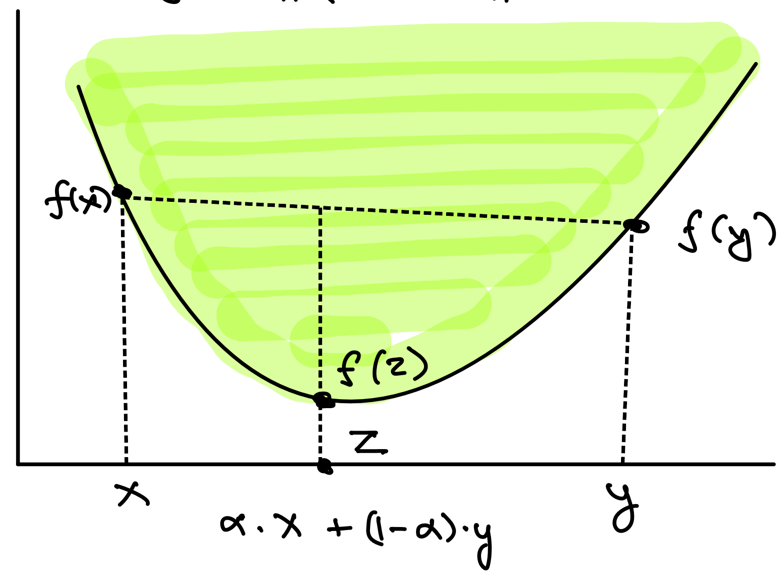


convex combination

$$u, v \in S \Rightarrow \alpha \cdot u + (1-\alpha)\cdot v \in S$$

$$\forall \alpha \in [0,1]$$

$f : \mathbb{R}^d \to \mathbb{R}$



$$f(z) \leq \alpha \cdot f(x) + (1-\alpha)\cdot f(y)$$

$$\downarrow$$

$$f(\alpha \cdot x + (1-\alpha)\cdot y)$$

g is concave $\Leftrightarrow$ $-g$ is convex

Jensen's inequality

Convex set S , $f : S \longrightarrow \mathbb{R}$

random variable X
supported in S

f is convex $\Rightarrow$ $\mathbb{E}[f(X)] \geq f(\mathbb{E}[X])$

f is concave $\Rightarrow$ $\mathbb{E}[f(X)] \leq f(\mathbb{E}[X])$

Exercises

- Prove Jensen's inequality when support of X is finite (induction)

- Prove $f(x) = x^2$, $f(x) = x \log x$ $\rightarrow$ convex

  $f(x) = \log x$ $\rightarrow$ concave

- Prove Cauchy-Schwarz using Jensen's

$$|\langle u, v \rangle| \leq \|u\|_2 \cdot \|v\|_2$$

Entropy

$$X = \{ a_1, \ldots \ldots, a_8 \}$$

$$p_1 \ldots \ldots \quad p_8$$

bits needed to provide certainty (specify one value)

$= ?$

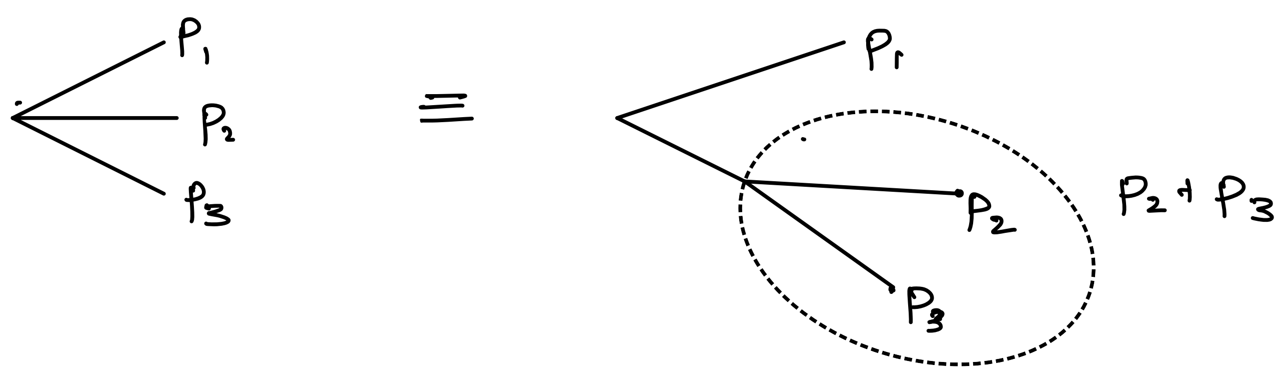
Entropy

$$\mathcal{X} = \{a_1, \ldots\ldots, a_n\}$$

$$p_1 \;\ldots\ldots\ldots\; p_n$$

$$\mathbb{P}[X = a_i] = p_i$$

- Continuous function of $p_1, \ldots\ldots, p_n$
- When $p_1 = \cdots = p_n = 1/n$. increasing in n
- Bundling:



Entropy

$$\mathbb{P}[X = a_i] = p_i \qquad H(X) = \sum_{i=1}^{n} p_i \cdot \log \frac{1}{p_i} \qquad \log_2$$

$$0 \log \frac{1}{0} = 0$$

$$p_1 = p_2 = \cdots = p_n = \frac{1}{n} \qquad H(X) = \sum_{i=1}^{n} \frac{1}{n} \log n = \log n$$

- $0 \le H(X) \le \log |\mathcal{X}| = \log n$

<u>Proof:</u> $p_1 \ldots \ldots p_n$

$$H(X) = \sum \underbrace{p_i}_{\geq 0} \cdot \underbrace{\log \frac{1}{p_i}}_{\geq 0}$$

$$H(X) = \sum_i p_i \cdot \log \frac{1}{p_i}$$

$$= \mathbb{E}\left[\log Y\right] \quad \text{for } Y = \frac{1}{p_i} \text{ w.p. } p_i$$

(Jensen's)

$$\le \log(\mathbb{E}\, Y)$$

$$= \log\left(\sum p_i \cdot \frac{1}{p_i}\right) = \log n$$

Source Coding

Prefix free code $C : \mathcal{X} \longrightarrow \Sigma^*$ s.t. $\forall\, x \neq y$

finite alphabet

$C(x) \neq C(y) \circ \sigma$

for any $\sigma \in \Sigma^*$

Fix $\Sigma = \{0,1\}$ (say)

Goal: Prefix-free C with least expected length

$$\mathbb{E}\,|C(X)| = \mathop{\mathbb{E}}_{x \sim P}\left[\,|C(x)|\,\right]$$

Kraft's inequality

Let $|X| = n$. $\exists$ prefix-free $C : X \to \{0,1\}^*$ with lengths $\ell_1, \ldots \ldots \ell_n$

if and only if $\displaystyle\sum_{i=1}^{n} \frac{1}{2^{\ell_i}} \leq 1$

"if" given $\ell_1 \ldots \ell_n$ s.t. $\sum 1/2^{\ell_i} \leq 1$ $\exists$ code

"only if" for all prefix-free C $\sum 1/2^{\ell_i} \leq 1$

(Assuming Kraft's inequality)

- Let C be any prefix-free code $C : \mathcal{X} \to \{0,1\}^*$.

  Then $\mathbb{E}\,|C(X)| \geq H(X)$

<u>Proof</u>:

$$H(X) - \mathbb{E}\,|C(X)|$$

$$= \sum p_i \log \frac{1}{p_i} - \sum p_i\, l_i$$

$$= \sum p_i \log\left(\frac{1}{p_i \cdot 2^{l_i}}\right)$$

$$= \mathbb{E}\left[\log Y\right] \qquad \underline{Y} = \frac{1}{p_i \cdot 2^{l_i}} \text{ w.p. } p_i$$

$$\leq \log(\mathbb{E}Y) = \log\left(\sum \cancel{p_i} \cdot \frac{1}{\cancel{p_i} \cdot 2^{l_i}}\right) = \log\left(\sum \frac{1}{2^{l_i}}\right) \leq \log 1 = 0$$

Shannon code

- There exists prefix-free $C : \mathcal{X} \to \{0,1\}^*$

  with $\mathbb{E}\,|C(X)| \leq H(X) + 1$

<u>Proof:</u> Construct code with $l_i = \left\lceil \log \frac{1}{p_i} \right\rceil$

$l_i \leq \log \frac{1}{p_i} + 1$

$\therefore \mathbb{E}\,|C(X)| = \sum p_i l_i \leq H(X) + 1$

$$\sum \frac{1}{2^{l_i}}$$

$$= \sum \frac{1}{2^{\lceil \log 1/p_i \rceil}}$$

$$\leq \sum p_i \leq 1$$

$\therefore \exists$ prefix-free code

Kraft's inequality

Let $|\mathcal{X}| = n$. $\exists$ prefix-free $C : \mathcal{X} \to \{0,1\}^*$ with lengths $\ell_1 \ldots \ldots \ell_n$

if and only if $\displaystyle\sum_{i=1}^{n} \frac{1}{2^{\ell_i}} \leq 1$

"if" given $\ell_1 \ldots \ell_n$ s.t. $\sum 1/2^{\ell_i} \leq 1$ $\exists$ code

"only if" for all prefix-free C $\sum 1/2^{\ell_i} \leq 1$

"only if" for all prefix-free C $\sum 1/2^{l_i} \leq 1$

Proof: Bingo!

Bingo tickets with codewords $C(x_1) \ldots C(x_n)$

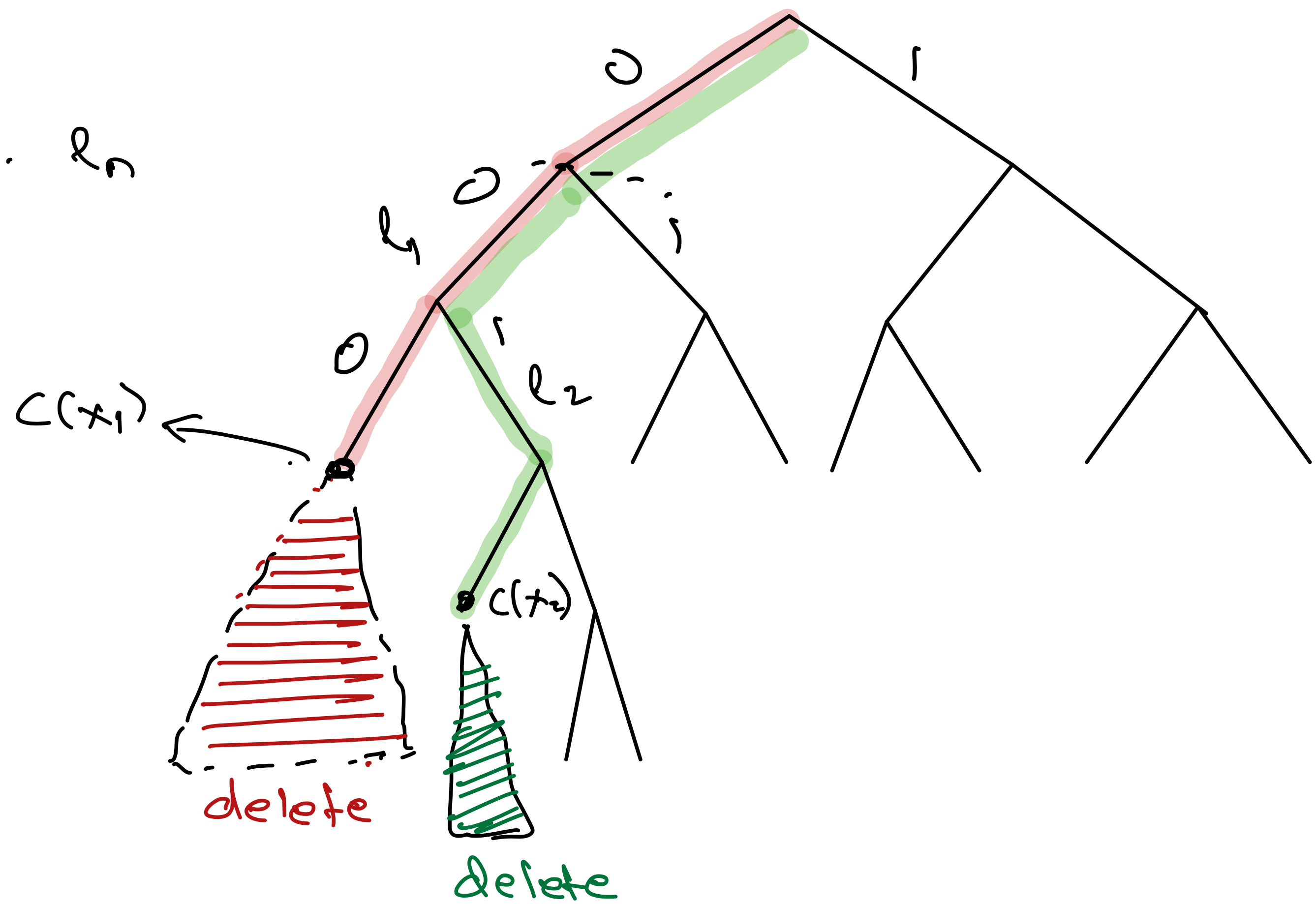
$\mathbb{P}[i^{th} \text{ ticket wins}] = \frac{1}{2^{l_i}}$ (need prefix-free)

disjoint events

$\therefore \sum \frac{1}{2^{l_i}} \leq 1$

"if" given $\ell_1 \dots \ell_n$ s.t. $\sum \frac{1}{2^{\ell_i}} \leq 1$ $\exists$ code

Sort $\ell_1 \dots \dots \ell_n$



- Can always pick next codeword if leaves left
- Deleted fraction of leaves $\frac{1}{2^{\ell_1}} + \dots + \frac{1}{2^{\ell_{i-1}}} < 1 \ \forall i$