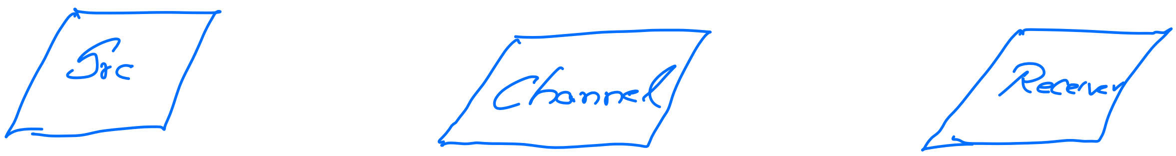
Today

- Shannon's Coding
  Noisy Coding Thm & Converse.
- BSC, BEC Channels

CS3.318.1
Coding Theory
Lecture 3 (2022-9-5)
Instructor: Prahladh Harsha.

Shannon:



When is communication feasible?

Mathematical Modelling: } via randomized means

Source : Distribution P, $x \sim P$

Rate : Entropy

Channel: $\mathcal{X} \ni x \rightarrow \square \rightarrow y \in \mathcal{Y}$

$\mathcal{X}$ - input } alphabets
$\mathcal{Y}$ - output

$\forall x \in \mathcal{X}$, there is a dist D_x on $\mathcal{Y}$

$x \rightarrow \square \rightarrow y$ $D_x(y) \cdot P(x)$

Memoryless Assumption

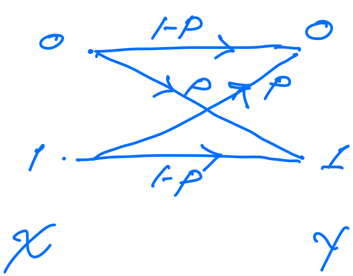
$$P(y_1 \ldots y_n \mid x_1 \ldots x_n) = \prod_{i=1}^{n} D_{x_i}(y_i)$$

Examples of Channels:

1. **Binary Symmetric Channel (p)** $[BSC_p]$

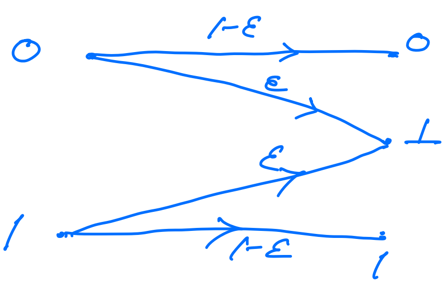
$p \in (0, \tfrac{1}{2})$

$\mathcal{X} = \mathcal{Y} = \{0, 1\}$



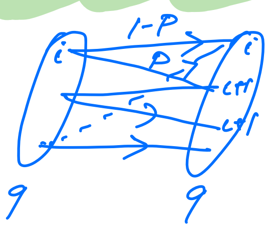
2. **Binary Erasure Channel (ε)** $[BEC_\varepsilon]$

$\mathcal{X} = \{0, 1\}$; $\mathcal{Y} = \{0, 1, \perp\}$



3. **Noisy-Typewriter Channel**

$\mathcal{X} = \mathcal{Y} = [9]$



(4) Binary Input w/ Additive Gaussian Noise

($BIAGN_\sigma$)

$X = \{1, -1\}$

$Y = \mathbb{R}$

$$P(y|x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{|x-y|^2}{2\sigma^2}}$$

Meta-Theorem:

$\forall$ memoryless discrete channel $\exists$ a number C - capacity of the channel.

(1) If Rate of source R satisfies $R < C$, then $\exists$ Enc. Dec

$$\text{Src} \xrightarrow{m} \text{Enc} \xrightarrow{E(m)} \text{Channel} \xrightarrow{y} \text{Dec} \xrightarrow{m'} \text{Receiver}$$

$y \sim D_{E(m)}$

$Pr[\text{error}]$ - Empy.

(2) If $R > C$, then $\forall$ Enc, Dec

$Pr[\text{error}] \approx 1$.

Noiseless Coding Theorem

: Channel is noiseless ie, identity function

Source S

Coding. When is compression feasible?

Noisy Coding Theorem: Source is incompressible (ie, uniform dist on source)

Channel Coding When is transmission feasible despite noisy channel?

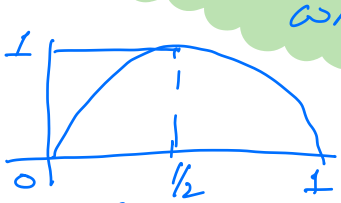
Meta Thm: obtained by noisy + noiseless

For the remaining lecture, focus on noisy coding theorem for BSC_p.

Capacity of BSC_p: $1 - H_2(p)$

where $H_2(p) = p \log_2 \frac{1}{p} + (1-p) \log_2 \left(\frac{1}{1-p}\right)$

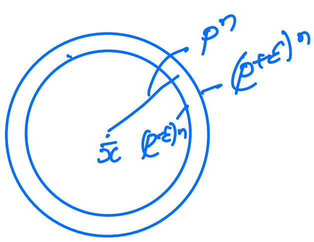
Binary entropy function.



Why does $h_2(p)$ occur?

$$x_1 \ldots x_n \xrightarrow{BSC_p} x_1 + e_1, x_2 + e_2 \ldots x_n + e_n$$

$$e_i \sim Ber(p)$$



$Vol_2(n, pn)$

Claim: $\mathrm{Vol}_2(n, pn) \leq 2^{h(p)n}$ for $p \in (0, 1/2)$

Pf: $$\frac{\mathrm{Vol}_2(n,pn)}{2^{h(p)n}} = \sum_{j=0}^{pn} \binom{n}{j} \cdot p^{pn}(1-p)^{(1-p)n}$$

$$= (1-p)^n \sum_{j=0}^{pn} \binom{n}{j} \left(\frac{p}{1-p}\right)^{pn}$$

$$\leq (1-p)^n \sum_{j=0}^{n} \binom{n}{j} \left(\frac{p}{1-p}\right)^{j} \quad \left[\text{Since } p < 1/2\right]$$

$$= (1-p)^n \left(1 + \frac{p}{1-p}\right)^n = 1$$

$\boxtimes$

Claim: $\mathrm{Vol}_2(n, pn) \geq 2^{h(p)n - o(n)}$

Pf: $$\mathrm{Vol}_2(n,pn) \geq \binom{n}{pn} \geq 2^{h(p)n - o(n)}$$

Stirling ~~for~~ approximation for $m!$

crude: $$\frac{m^m}{e^{m-1}} \leq m! \leq \frac{m^{m+1}}{e^{m-1}}$$ (obtained from $\sum_{i=1}^{m-1} \ln i \leq \int_1^m \ln x \leq \sum_{i=2}^{m} \ln i$)

$\boxtimes$

Noisy Coding Theorem

$\forall p \in (0, 1/2)$, $\varepsilon \in (0, \frac{1}{2} - p)$

$\exists\, \delta, n_0 \quad \forall n \geq n_0$

$\exists\, k = \lfloor (1 - H(p+\varepsilon))n \rfloor$

s.t. $E: \{0,1\}^k \to \{0,1\}^n$ &

$D: \{0,1\}^n \to \{0,1\}^k \cup \{\perp\}$

$$\Pr_{\substack{m \leftarrow_U \{0,1\}^k \\ e \leftarrow (\mathrm{Ber}(p))^n}} [D(E(m)+e) \neq m] \leq 2^{-\delta n}$$

Proof: What is a suitable encoding function?

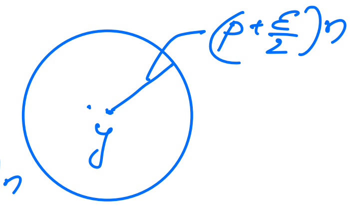
- Pick E probabilistically.

  For each $m \in \{0,1\}^k$, pick $E(m)$ - random n-bit string (independently).

- What is decoding function D?

  $y \in \{0,1\}^n$

  $D(y)$ = If $\exists$ a unique m s.t. $\Delta(E(m), y) < (p+\frac{\varepsilon}{2})n$
  return m
  else return $\perp$.



Bad events:

$E1$: $\sum e_i \geq (p+\frac{\varepsilon}{2})n$
(ie too many errors)

$E2$: $\exists\, m' \neq m$, st $\Delta(E(m'), E(m)+e) < (p+\frac{\varepsilon}{2})n$
(ie, m & e are a bad message, noise pair.)

Obs: $\neg E_1 \wedge \neg E_2 \Rightarrow$ Decoding is correct

(i.e., $D(E(m)+e) = m$.)

Hence it suffices to bound $\Pr[E_1 \vee E_2]$

$\Pr[E_1] \leq \exp(-C\varepsilon^2 n)$ (via Chernoff)

$$\Pr[E_2] = \Pr_{E,e}\left[\exists m' \neq m,\ \Delta(E(m'), E(m)+e) < (p+\tfrac{\varepsilon}{2})n\right]$$

Fix $E(m)$, e, $m' \neq m$

$$\Pr_{E(m')}\left[\Delta(E(m'), E(m)+e) < (p+\tfrac{\varepsilon}{2})n\right]$$

$$= \frac{\mathrm{Vol}_2(n, (p+\frac{\varepsilon}{2})n)}{2^n}$$

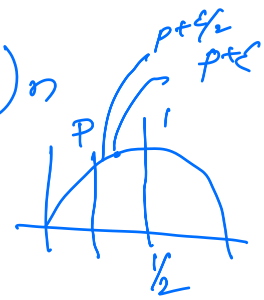
$$\leq 2^{-n(1-h(p+\frac{\varepsilon}{2}))}$$

$$\Pr_{E(m')(m'\neq m)}\left[\exists m',\ \Delta(E(m'), E(m)+e) < (p+\tfrac{\varepsilon}{2})n\right]$$

$$\leq 2^k \cdot 2^{-n(1-h(p+\frac{\varepsilon}{2}))}$$

$$= 2^{(h(p+\frac{\varepsilon}{2}) - h(p+\varepsilon))n}$$

$$\leq 2^{-f(p,\varepsilon)n}$$



Choose δ s.t.

$$2^{-C\varepsilon^2 n} + 2^{-f(p,\varepsilon)n} \leq 2^{-\delta n}.$$

$\boxtimes$

Question: Can one do better than $1-h_2(p)$?

No!

Converse Coding Theorem

$\forall p \in (0, \frac{1}{2})$, $\varepsilon \in (0, \frac{1}{2}-p)$

$\exists \delta$, n_0, $\forall n \geq n_0$

if $k \geq (1-h_2(p)+\varepsilon)n$

$\forall$ E, D - functions

$\exists m$, $\Pr_{e \leftarrow (Ber(p))^n}[D(E(m)+e)=m] \leq 2^{-\delta n}$ ... (*)

Pf: (Intuition).

Suppose (*) was false

then for each m, $\exists$ $T_m \subseteq \{0,1\}^n$

s.t. ① $D(T_m) = m$

② $|T_m|$ - large

$\Rightarrow 2^k$ large $< 2^n$

contradiction

(formal proof next time).