Channel Coding Theorem

* The most fundamental theorem in inf. theory.
* Arguably the first application of the probabilistic method in math.

We suppose the message $W$ is drawn from an index set $\{1, \ldots, M\}$, uniformly at random (turns out to be not different from worst-case).

- Encoder $X^n_i : \{1, \ldots, M\} \rightarrow \mathbb{X}^n$.

  $\text{Supp} (X^n_i) = \text{"Code book"}$.

  Each possible output is a "codeword".

- Decoder: Deterministic function

  $g: \mathbb{Y}^n \rightarrow \{1, \ldots, M\}$

- Rate: $R = \frac{\log_2 M}{n}$ bits/channel use.

* Error probability:

  Fix $i \in \{1, \ldots, M\}$, $X_i := \Pr (g(Y^n) \neq i \mid W = i)$.

  $P_e = \frac{1}{M} \sum_{i=1}^{M} X_i = \Pr_{(\text{uniform } W)} (g(Y^n) \neq W)$. 
* Def: "Achievable rate" $R$ is achievable if

\[ P_e \to 0 \text{ as } n \to \infty. \]

\[ \exists \text{ sequence of codes at rate } \geq R \text{ s.t.} \]

* Def: $R^* \triangleq \sup \{ R \mid R \text{ is achievable} \}$.

Channel Coding Theorem: $R^* = C$. 

Achievable
* Achievability Proof: *

[Existence]

Fix some $R < C$ and $p(x)$.

[We create a random codebook and show it performs well.]

Codebook, $C = \begin{pmatrix} X_1(1) & \cdots & X_n(1) \\
                    X_1(2) & \cdots & X_n(2) \\
                    \vdots  & \ddots & \vdots  \\
                    X_1(R^n) & \cdots & X_n(R^n) \end{pmatrix}$

* Each entry is iid according to $p(x)$.

[Sender and receiver use this code.]

* Suppose a uniform random $W \in \{1, \ldots, M\}$ is sent.

* Receiver uses "jointly typical decoding": Given $Y^n$, find $X^n$ such that

  $(X^n, Y^n)$ is jointly typical.

  If $X^n$ is unique, decode to the corresponding index.

  Else, declare an error (output $0^n$).
\[ \text{def's } E: \text{ error event} \]

\[ \Pr(E) = \sum_{\mathcal{E}} \Pr(C) \cdot P_E^{(n)}(C) \]

\[ = \sum_{\mathcal{E}} \Pr(C) \cdot \frac{1}{2^n R} \sum_{w=1}^{2^n R} \lambda_w(C). \quad (1) \]

* Suppose now that \( W \) is fixed to \( W=1 \).

* def: \( E_i \): event that the \( i \)-th codeword \( \tilde{y}_i \) is jointly typical with \( y^n \).

\[ \Rightarrow \Pr(E | W=1) \leq \Pr(\bigcup \neg E_i | W=1) + \]

\[ \sum_{i=2}^{2^n R} \Pr(E_i | W=1). \]

* Take \( n \) large so that \( \Pr(\neg E_1 | W=1) \leq \varepsilon \) (by AEP)

* Also \( X^n_j(1) \) and \( X^n_i(i) \) for \( i \neq 1 \) are independent,

so by AEP, thus \( X^n_j(1) \perp X^n \)

by AEP, \( \Pr(E_i | W=1) \leq 2^{-n(I(X^n_j;Y)-3\varepsilon)} \)
\[ \Pr(\mathcal{E}|W=1) \leq \varepsilon + \sum_{i=2}^{nR} 2^{-n(I(X;Y)-3\varepsilon)} \leq \varepsilon + 2^{-n(I(X;Y)-R)} \leq 2\varepsilon, \]
if \( R < I(X;Y) - 3\varepsilon \) and \( n \) large.

\[ \text{Since this} \]
\[ \text{True for all fixings of } W \Rightarrow \Pr(\mathcal{E}) \leq 2\varepsilon. \]

\[ \text{Now take } p(x) \text{ to be } p^*(x) \text{ in capacity} \]
\[ \Rightarrow R < C \text{ is achievable.} \]

\[ \text{Fix the code. Can be found by exhaustive search.} \]

\[ \text{Get worst case small } \Pr(\mathcal{E}). \]
\[ (i.e., \text{for all } W). \]

\[ \text{Markov: } \Pr(W \mid X \geq 4\varepsilon) \leq \frac{1}{2} \]

\[ \text{Throw away half of the codewords} \]
\[ \Rightarrow \text{New rate } = R - \frac{1}{n} \]
\[ \frac{2^{-nk=1 \text{ codewords}}}{\text{[2}^{-nk=1 \text{ codewords]}}} \]