Almose loseless / errorless. Coding

Coding scheme:

$$\mathcal{X} \xrightarrow[f]{\text{encoder}} \{0,1\}^* \xrightarrow[g]{\text{decoder}} \mathcal{X}$$

Idea: If we want perfect (errorless) code, then $k \geq \log_2 |\mathcal{X}|$

But any phsysic process is subjected to noise,

the transmission is erroneous.

In reality, we do things with certain accuracy / error tolerance.

Def: A pair $(\underset{\text{encoder}}{f}, \underset{\text{decoder}}{g})$ is called a (k, ε)-code (fixed length) for X if

$$f: \mathcal{X} \to \{0,1\}^k \qquad g: \{0,1\}^k \to \mathcal{X}$$

k bits string.

Such that $P(g(f(X)) \neq X) \leq \varepsilon$

undectable error

Remark: Alternative setting with detectable error

$$f: \mathcal{X} \to \{0,1\}^k \qquad g: \{0,1\}^k \to \mathcal{X} \cup \{e\}$$

such that $\exists S \subseteq \mathcal{X}$ s.t. $g(f(x)) = \begin{cases} x & x \in S \\ e & x \notin S \end{cases}$

$$P(g(f(X)) \neq X) = P(g(f(X)) = e) = P_X(S)$$

S lossless part. $\quad e$ detectable error

Def: (Optimal error probability)

$$\varepsilon^*(X,k) := \inf\{\varepsilon : \exists (k,\varepsilon)\text{-code for } X\}$$

Smallese error that a longeh k code can acheive.

Thm: $\varepsilon^*(X,k) = P[l(f^*(X)) \geq k] = 1 - \sum_{i=1}^{2^k-1} P_X(i)$

(Recall that we asume $P_X(i) \geq P_X(i+1)$) decreasing

Pf: We assign a code word to each of $\underbrace{2^k-1 \text{ most likely realization of } X}_{S}$.
and all the rese to one word "error"

$$\varepsilon^*(X,k) = P[X \notin S] = P[l(f^*(X)) \geq k]$$

Actual code book: Variable longeh $\{\emptyset, 0, 1, 00, 01, \cdots \underbrace{11\cdots1}_{k-1}\}$

fixed length $\{(00\cdots00)\quad(00\cdots01)\quad\cdots\quad\underbrace{(11\cdots10)}_{2^k-1}\}$

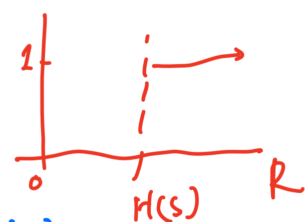
$(11\cdots11)$ for "error". $\square$

Shannon's Sorrce coding Thm 1948

Let S^n be i.i.d. Then

Rate

$$\lim_{n\to\infty} \varepsilon^*(S^n, nR) = \begin{cases} 0 & \text{if } R > H(S) \\ 1 & \text{if } R < H(S) \end{cases}$$



Lemma (Achievability)

$$\varepsilon^*(X,k) \leq P\left[\log_2 \frac{1}{P_X(X)} \geq k\right]$$

Indeed, $\varepsilon^*(X,k) = \sum_{m \geq 2^k} P_X(m) = \sum 1_{\{m \geq 2^k\}} P_X(m) \leq \sum 1_{\{\frac{1}{P_X(m)} \geq 2^k\}} P_X(m)$

$$= \mathbb{E}\, 1_{\{\log_2 \frac{1}{P_X(X)} \geq k\}}$$

Lemma (converse)

$$\varepsilon^*(X,k) \geq P\left[\log \frac{1}{P_X(X)} > k+\tau\right] - 2^{-\tau} \quad . \quad \forall \tau > 0$$

Denote, $L = l \circ f^*$. $\quad L(m) = \lfloor \log m \rfloor \leq \log m \leq \log \frac{1}{P_X(m)}$

$$1 - \varepsilon^*(X,k) = \quad P[L \leq k]$$

$$= P\left[L \leq k, \ \log_2 \frac{1}{P_X} \leq k+\tau\right] + P\left[L \leq k, \log_2 \frac{1}{P_X} > k+\tau\right]$$

$$\leq P\left[\log_2 \frac{1}{P_X} \leq k+\tau\right] + \sum_m P_X(m) \, 1_{\{L(m) \leq k\}} 1_{\{P_X(m) \leq 2^{-k-\tau}\}}$$

$$\leq P\left[\log_2 \frac{1}{P_X} \leq k+\tau\right] + (2^{k+1}-1) \cdot 2^{-k-\tau}$$

(Pf of Thm)

$$P\left[\log_2 \frac{1}{P_X} > k+\tau\right] - 2^{-\tau} \leq \varepsilon^*(X,k) \leq P\left[\log \frac{1}{P_X} \geq k\right]$$

Now take $X = S^n$, $\quad P_{S^n}(s_1 \cdots s_n) = P_S(s_1) \cdot P_S(s_2) \cdots P_S(s_n)$

By WLLN, $\frac{1}{n} \log \frac{1}{P_{S^n}} = \frac{1}{n}\left(\log \frac{1}{P_{S_1}} + \log \frac{1}{P_{S_2}} + \cdots \log \frac{1}{P_{S_n}}\right)$

$$\xrightarrow{P} \quad \mathbb{E}\left(\log \frac{1}{P_S}\right) = H(S)$$

Then

$$\varepsilon^*(S^n, nR) \leq P\left[\log \frac{1}{S^n} > nR\right]$$

$$= P\left[\frac{1}{n} \log \frac{1}{S^n} > R\right] \longrightarrow P\left[\underset{\overset{\downarrow}{\text{constant}}}{H(S)} > R\right] = 0$$

if $R > H(S)$

$$\varepsilon^*(S^n, nR) \geq P\left[\log \frac{1}{P_{S^n}} > nR - \sqrt{n}\right] - 2^{-\sqrt{n}}$$

$$\geq P\left[\frac{1}{n} \log \frac{1}{P_{S^n}} > R - \frac{1}{\sqrt{n}}\right] - 2^{-\sqrt{n}}$$

$$\geq P\left[\frac{1}{n} \log \frac{1}{P_{S^n}} > R\right] - 2^{-\sqrt{n}}$$

$$\lim_{n\to\infty} \varepsilon^*(S^n, nR) \geq \lim_{n\to\infty} P\left[\frac{1}{n}\log\frac{1}{P_{S^n}} > R\right] - \lim_{n\to\infty} 2^{-\sqrt{n}}$$

$$\geq \quad P[H(S) > R] = 1$$

if $H(S) > R$.

A Direct argument. Typical Sequence

Let $\mathcal{X}$ be a finite alphabet.

For a sequence $x^n = x_1 \cdots x_n \in \mathcal{X}^n$, we can define

$$N_{x^n}(a) = \#\{i \mid x_i = a\}$$

Then $\frac{N_{x^n}(a)}{n}$ is the freqnency "a" appears in x^n

$P_{x^n}(a) := \frac{N_{x^n}(a)}{n}$ emperical distrinbition of x^n.

Now

$X \sim P$ R.V. on $\mathcal{X}$, $X^n = X_1 \cdots X_n$ i.i.d

X^n is a random sequence.

What would be mosely likely sequence?

$$P_{X^n} \approx P_X$$

A sequence x^n is strong (n,ε)-typical for P if

$$|P_{x^n}(a) - P_X(a)| < \frac{\varepsilon}{|\mathcal{X}|} \quad \forall a \in \mathcal{X}$$

is weak (n,ε)-typical if

$$H(P)+\varepsilon < \frac{1}{n}\log\frac{1}{P_{X^n}(x^n)} < H(P)-\varepsilon$$

We denote $T_\varepsilon^n(P) \subseteq \mathcal{X}^n$ be the set of strong typical sequence of P

$A_\varepsilon^n(P) \subseteq \mathcal{X}^n$ - - - weak - - - - of P

$\boxed{\begin{array}{l} \text{Theorem (Typicality) ① } T_\varepsilon^n(P) \subseteq A_\varepsilon^n(P) \\ \text{② } (1-\varepsilon)2^{n(H(X)-\varepsilon)} \leq |A_\varepsilon^n(P)| \leq 2^{n(H(P)+\varepsilon)} \\ \text{③ } 2^{n(H(X)-\delta(\varepsilon))} |T_\varepsilon^n(P)| \leq 2^{n(H(P)+\delta(\varepsilon))} \\ \quad \text{where } \delta(\varepsilon) \to 0 \text{ as } \varepsilon \to 0 \\ \text{④ } P_{X^n}(x^n \in A_\varepsilon^n) \to 1 \\ \quad P_{X^n}(X^n \in T_\varepsilon^n) \to 1 \text{ as } n \to \infty \end{array}}$

$$x^n \in A_\varepsilon^n(P) \Leftrightarrow 2^{-n(H(P)+\varepsilon)} < P_{X^n}(x^n) < 2^{-n(H(P)-\varepsilon)}$$

$$P_{X^n}(X^n \in A_\varepsilon^n) = P_{X^n}\left(\left|\frac{1}{n}\sum_{i=1}^{n}\log\frac{1}{P_{X_i}} - H(X)\right| \leq \varepsilon\right) \to 1$$

by W.L.L.N

Alternative proof for Shannon's Coding Theorem.

Choose error free set $S(n) = A_\varepsilon^n$. Assign a codeword to each $x^n \in S(n)$.

need $|S(n)| \leq 2^{n(H(X)+\varepsilon)}$ codeword

word length $\leq \lfloor n(H(X)+\varepsilon) \rfloor$

error probality $\varepsilon(n) = 1 - P(S(n)) \longrightarrow 0$

$\square$