ST06 Lecture 14

Note Title 4/4/2006

Today:

Error Exponents in BSC decoding

First: a review of course so far...

Phase I: Entropy, Mutual Information, ...

(Tools to be used later)

Phase II: Source Coding

"How to compress source nicely"

AEP, Shannon Coding, Kraft's Inequality

Huffman, Lempel-Ziv Coding.

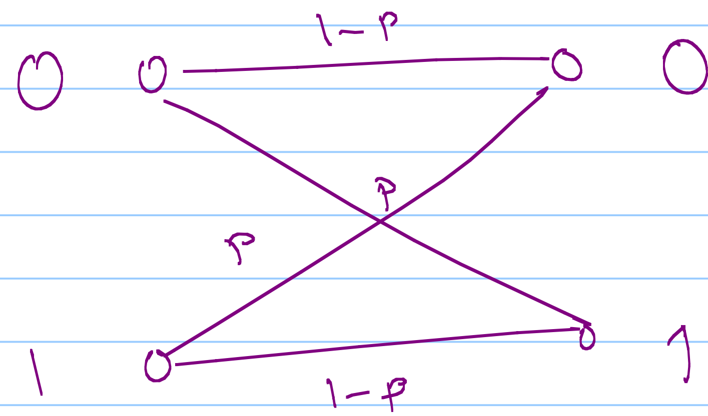
Phase III: Channel Coding:

"How to protect against channel infidelity?"

Channel Capacity, Random Coding, MLD Decoding, Joint AEP, Coding Theorem Converse.

Today: Closer look at the BSC.

What?



Know: Capacity $= 1 - H(p)$

where $H(p) = p \log \frac{1}{p} + (1-p) \log \frac{1}{1-p}$.

Implies: can map $E: Rn$ bits $\to$ n bits

s.t. $P_{err} \triangleq \Pr\left[D(\text{Channel}(E(x))) \neq x \right] \to 0$

provided $R < C$.

So how does P_{err} grow with R, C, n?

Easy to see $P_{err} = \exp(-n)$ if $R < C$

Question: if $P_{err} = 2^{-E_c(R) \cdot n}$ then

what is $E_c(R)$?

Why?

Recall: Channel Coding is about "improving" reliability (not making channel error-free).

So how much does reliability improve?

Phase 0 :

Why exponential error ?

Observation 1 :

if Code has two codewords, then

$$P_{err} \geq c^{-n} \qquad [\geq p^n !]$$

Observation 2 :

Recall random coding + analysis

- Code picks $E(1) \ldots E(2^{Rn})$ at random from $\{0,1\}^n$.

- Decode 1 :

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

m s.t. $\Pr[y \mid E(m)]$ is maximum.

- Decode 2:
  - Pick threshold $T \in [p, H^{-1}(1-R)]$
  - if $\exists!\ m$ s.t.

$$\Delta(y, E(m)) \leq T \cdot n$$

then output m, else output ERROR.

Analysis (of Decode 2)

Error (type 1):

transmit $E(m)$ but $\Delta(y, E(m)) > T \cdot n$

$$\Pr[\text{type 1 error}] \longrightarrow 0 \quad \text{if } T > p.$$

[Chernoff bound]...

Error (type 2):

$\Delta(E(m'), y) \leq \tau \cdot n$ for some $m' \neq m$

$\Pr[\text{type 2 error}] \longrightarrow 0$

[provided $H(\tau) + R < 1$]

Question: How fast do these quantities go to zero?

$\Pr[\text{type 1}]$

$$= \sum_{i=\tau n}^{n} \binom{n}{i} p^{i} (1-p)^{n-i}$$

$$\approx \binom{n}{\tau n} \cdot p^{\tau n} (1-p)^{(1-\tau)n}$$

$$\approx 2^{n\left[\tau \log \frac{1}{\tau} + (1-\tau)\log\frac{1}{1-\tau} + \tau \log p + (1-\tau)\log(1-p)\right]}$$

$$= 2^{-n\left[\tau \log \frac{\tau}{p} + (1-\tau)\log\frac{(1-\tau)}{p}\right]}$$

$$= 2^{-D(\tau \| p)\cdot n}$$

[Abusing notation
$\tau \equiv \mathrm{Bern}(\tau)$
$p \equiv \mathrm{Bern}(p)$]

$$\boxed{\Pr[\text{type I error}] \approx 2^{-D(\tau\|p)\cdot n}}$$

[Note: has nothing to do with code;
But element only of our analysis]

Type II

Pr[type 2 error]

$$= 1 - \left(1 - \frac{\sum_{i=0}^{\tau n} \binom{n}{i}}{2^n}\right)^{2^{Rn} - 1}$$

[Only for random code!]

$$\approx 1 - \left(1 - 2^{(H(\tau) - 1) n}\right)^{2^{Rn}}$$

$$\approx 1 - \left(1 - 2^{Rn} \cdot 2^{(H(\tau) - 1) n}\right)$$

$$\approx 2^{(R + H(\tau) - 1) n}$$

Putting things together

Error occurs if $\exists\, \tau$ such that

$$\Delta(y, E(m)) > \tau \cdot n \qquad (\text{Type I}, \tau)$$

AND $\exists\, m'$ s.t. $\Delta(y, E(m')) \leq \tau n$

$$(\text{Type II}, \tau)$$

$$\max_{\tau} \begin{bmatrix} \text{Type 1 error} \\ \text{\& Type 2 error} \end{bmatrix} \leq P_{err} \leq \underset{\tau}{\text{Sum}} \begin{bmatrix} \text{Type I error} \\ \text{\& Type II error} \end{bmatrix}$$

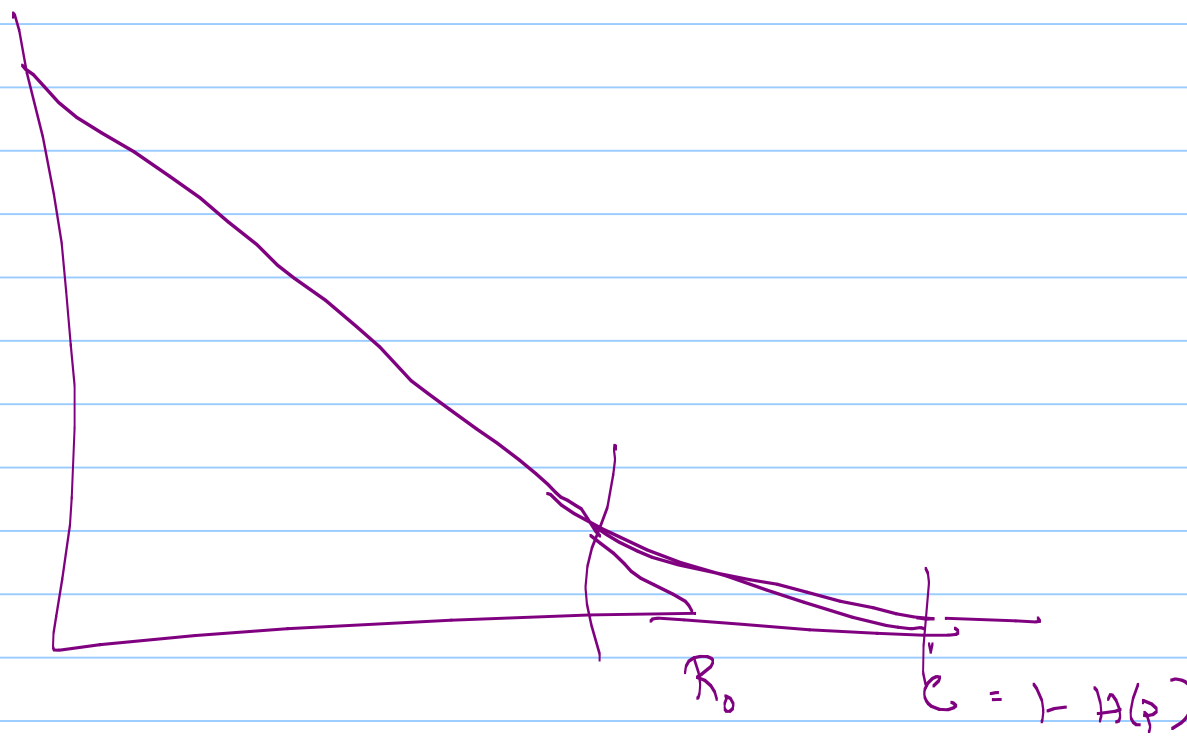
[within factor of $(n+1)$ of each other]

$$P_{err} \approx \max_{\tau} \left\{ 2^{-D(\tau \| p) n} \cdot 2^{(R + H(\tau) - 1) n} \right\}$$

Error Exponent [for random coding]

$$E_{RCE}(R) = \min_{p \le \tau \le H^{-1}(1-R)} \left\{ D(\tau \| p) + D(\tau \| \tfrac{1}{2}) - R \right\}$$

Plot of $E_{RCE}(R)$ for $p = .007$



So what does $E_{RCE}(R)$ look like?

linear in R for $R \leq R_{crit}$

convex for $R_{crit} \leq R \leq C$.

$R_{crit} = ?$

linear $= ?$

convex $= ?$

$R = 0$? want $D'(T \| p) = D'(T \| \frac{1}{2})$

optimization yields $\frac{T}{1-T} = \frac{\sqrt{p}}{\sqrt{1-p}}$

maximize $D(T\|p) + D(T\|\tfrac{1}{2})$

maximized at $T = H^{-1}(1-R)$

or at $\dfrac{T}{1-T} = \dfrac{\sqrt{p}}{\sqrt{1-p}}$ $[T_{crit}]$

[if $R \ll 1 - H(p)$ then T_{crit} is bottleneck

if $R \to 1 - H(p)$ then T_{crit} is not bottleneck]

if T_{crit} is bottleneck

then $E(R) = R_0 - R$ $\quad R_0 = D(T\|p) + D(T\|\tfrac{1}{2})$

$$R_0 = 1 - \log\left(1 + 2\sqrt{p(1-p)}\right)$$

if T_{crit}

$$E(R) = D(R \| p) + D(R \| \tfrac{1}{2}) - R$$

Observations

Above analysis focused on

1. Random Coding
2. "Typical Decoding"

$\Rightarrow$ Yields "lower bound" on error exponent.

But can reverse most steps for

"random code" to say it is the

<u>right</u> exponent for random codes.

But is it the right exponent for BSC?

(best code & best decoding?)

<u>NO</u>: Expurgated codes.

- Pick random codes
- Throw away some "bad words"
- achieve better exponent for small R.

What is the best exponent?

Unknown.

See

Barg & Forney

IEEE IT, vol. 48 no. 9,

Sept. 2002

for more