EE4-40 Information theory

Solutions 2015

[A] application
[B] book word
[T] new theory
[E] new example

1\. a) Just expand RHS

$$H(\mathbf{q}) + q_2 H\left(\frac{p_2}{q_2}, \frac{p_3}{q_2}\right)$$

$$= -q_1 \log q_1 - q_2 \log q_2 - \cancel{q_2}\frac{p_2}{\cancel{q_2}}\log\frac{p_2}{q_2} - \cancel{q_2}\frac{p_3}{\cancel{q_2}}\log\frac{p_3}{q_2}$$ [2E]

$$= -p_1\log p_1 - p_2\log p_2 - p_3\log p_3 \underbrace{+ p_2\log q_2 + p_3\log q_2 - q_2\log q_2}_{=0 \text{ since } q_2 = p_2+p_3}$$ [2E]

$$= -p_1\log p_1 - p_2\log p_2 - p_3\log p_3$$ [2E]

$$= H(\mathbf{p})$$

b) We have

$X_1 X_2$	00	01	10	11
Y	0	0	0	1

each with prob. 1/4 [1E]

i) Thus

$p(y=0) = \frac{3}{4}$ $\quad p(y=1) = \frac{1}{4}$

$$H(Y) = -\frac{1}{4}\log\frac{1}{4} - \frac{3}{4}\log\frac{3}{4}$$ [2E]

$$= \frac{1}{2} + 0.31$$

$$= 0.81$$

ii) We have the conditional distribution

$X_1 \backslash Y$	0	1
0	1	0
1	$\frac{1}{2}$	$\frac{1}{2}$

[3E]

Thus

$$H(Y|X_1) = \frac{1}{2}H(1) + \frac{1}{2}H\left(\frac{1}{2}\right) = \frac{1}{2}$$

$$I(X_1;Y) = H(Y) - H(Y|X_1)$$

$$= 0.81 - \frac{1}{2} = 0.31$$

iii) $I(X_{1:2};Y) = H(Y) - H(Y|X_{1:2})$

$= H(Y) - 0$ $\quad$ Y is a function of $X_{1:2}$

$= H(Y) = 0.81$ [3E]

c) X takes values in $\{1, 2, 3, \cdots\}$.

$X = n$ means that Tail occurs for the first $n-1$ flips, while Head occurs for the n-th flip. [3 A]

Thus

$$p(X=n) = \left(\tfrac{1}{2}\right)^{n-1}\tfrac{1}{2} = \left(\tfrac{1}{2}\right)^{n} \qquad [3\ A]$$

$$H(X) = \sum_{n=1}^{\infty} 2^{-n}\log 2^{n}$$

$$= \sum_{n=1}^{\infty} n \cdot 2^{-n} \cdot \log 2 \qquad [4\ A]$$

$$= \frac{\frac{1}{2}}{\left(1-\frac{1}{2}\right)^2}$$ use the second formula for $r = \frac{1}{2}$

$$= 2 \text{ bits}$$

2. a)

i) The marginal distributions are given by

$$p_X(x) = p_Y(y) = \{\tfrac{1}{4}, \tfrac{3}{4}\}$$ [1 E]

So their entropy is

$$H(X) = H(Y) = H(\tfrac{1}{4}, \tfrac{3}{4})$$ [2 E]

$$= -\tfrac{1}{4}\log\tfrac{1}{4} - \tfrac{3}{4}\log\tfrac{3}{4}$$

$$= 0.81$$

Since ~~the~~ for sequences $\mathbf{X}$ and $\mathbf{Y}$

$$-\tfrac{1}{8}\log p(\mathbf{X}) = -\tfrac{1}{8}\log p(\mathbf{Y})$$

$$= -\tfrac{1}{8}\log(\tfrac{1}{4})^2(\tfrac{3}{4})^6$$ [2E]

$$= -\tfrac{1}{4}\log\tfrac{1}{4} - \tfrac{3}{4}\log\tfrac{3}{4}$$

$$= H(X) = H(Y),$$

both of them are typical. [1 E]

The joint distribution is $p_{XY}(x,y) = \{\tfrac{1}{8}, \tfrac{1}{8}, \tfrac{1}{8}, \tfrac{5}{8}\}$

$$H(X,Y) = -\tfrac{3}{8}\log\tfrac{1}{8} - \tfrac{5}{8}\log\tfrac{5}{8} = 1.55$$ [1 E]

For sequence $(\mathbf{X}, \mathbf{Y})$, we check

$$-\tfrac{1}{8}\log p(\mathbf{X}, \mathbf{Y}) = -\tfrac{1}{8}\log(\tfrac{1}{8})^4(\tfrac{5}{8})^4$$ [2E]

$$= -\tfrac{1}{2}\log\tfrac{1}{8} - \tfrac{1}{2}\log\tfrac{5}{8}$$

$$= 1.84 > H(X,Y) + \varepsilon = 1.55 + 0.2$$

So they are not jointly typical. [1 E]

ii) (1) total probability theorem [1B]

(2) first term. taking maximum yields an upper bound

Second term: $P(A \cap B) \le P(B)$ [1 B]

(3) $|S^{(n)}| < 2^{n(H(X)-2\varepsilon)}$ is given

$p(\mathbf{X}) < 2^{-n(H-\varepsilon)}$ if typical [2B]

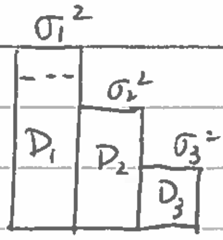
probability of atypical set $< \varepsilon$

(4) algebra

(5) $2^{-n\varepsilon} < \varepsilon$ because $n > -\varepsilon^{-1}\log\varepsilon$ [1 B]

b)

i) If only one source is encoded,



$3D = D_1 + D_2 + D_3$

$> \sigma_2^2 + \sigma_2^2 + \sigma_3^2$

$= 2\sigma_2^2 + \sigma_3^2 \Rightarrow D > \frac{2\sigma_2^2 + \sigma_3^2}{3}$

[3A]

$R_1 = \frac{1}{2}\log\frac{\sigma_1^2}{D_1}$

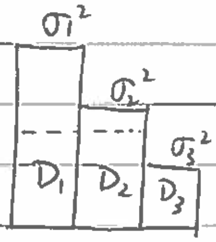
$D_1 = 3D - D_2 - D_3$

$= 3D - \sigma_2^2 - \sigma_3^2$

$= \frac{1}{2}\log\frac{\sigma_1^2}{3D - \sigma_2^2 - \sigma_3^2}$

$R_2 = R_3 = 0$

ii) If two sources are encoded



$3D = D_1 + D_2 + D_3$

$> 3\sigma_3^2$

$3D < 2\sigma_2^2 + \sigma_3^2$

[3A]

$\sigma_3^2 < D < \frac{2\sigma_2^2 + \sigma_3^2}{3}$

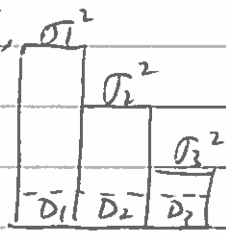
$D_1 = D_2 = \frac{3D - \sigma_3^2}{2}$

$R_1 = \frac{1}{2}\log\frac{\sigma_1^2}{D_1} = \frac{1}{2}\log\left(\frac{2\sigma_1^2}{3D - \sigma_3^2}\right)$

$R_2 = \frac{1}{2}\log\frac{\sigma_2^2}{D_2} = \frac{1}{2}\log\left(\frac{2\sigma_1^2}{3D - \sigma_3^2}\right)$

$R_3 = 0$

iii) If all three sources are encoded,



$D < \sigma_3^2$

$D_1 = D_2 = D_3 = D/3$

[3A]

$R_1 = \frac{1}{2}\log\frac{\sigma_1^2}{D_1} = \frac{1}{2}\log\frac{3\sigma_1^2}{D}$

$R_2 = \frac{1}{2}\log\frac{3\sigma_2^2}{D}$

$R_3 = \frac{1}{2}\log\frac{3\sigma_3^2}{D}$

3. a) (1) average over codewords, average over codes [1 B] each

(2) exchange order of summation

(3) for random coding, $\sum_C p(C)\lambda_w(C)$ doesn't depend on index w, so w can be 1 w.l.o.g.

(4) average error prob of $w=1$

(5) definition of error prob

(6) Union bound

(7) prob. of atypical set $\le \varepsilon$

prob. of $X(w)$ and y jointly typical $\le 2^{-n(I(X;Y)-3\varepsilon)}$

(8) algebra: $2^{nR}-1 < 2^{nR}$

(9) $I(X;Y) \le C$

(10) $n > -\frac{\log \varepsilon}{C-R-3\varepsilon} \Rightarrow 2^{-n(C-R-3\varepsilon)} < \varepsilon$

(11) by contradiction

(12) by contradiction

(13) Since half of the codewords are gone,

$$\text{rate} = \frac{1}{n}\log(2^{nR}/2) = \frac{nR-1}{n} = R - n^{-1}$$

b)

i)

$$\begin{aligned} Y &= \alpha Y_1 + (1-\alpha)Y_2 \\ &= \alpha(X+Z_1) + (1-\alpha)(X+Z_2) \\ &= X + \alpha Z_1 + (1-\alpha)Z_2 \end{aligned}$$ [3 A]

So this is an equivalent Gaussian channel with noise variance $\alpha^2 N_1 + (1-\alpha)^2 N_2$.

Capacity

$$C = \frac{1}{2}\log\left(1 + \frac{P}{\alpha^2 N_1 + (1-\alpha)^2 N_2}\right)$$ [3 A]

ii) To maximize C, we minimize noise variance

Let $\frac{\partial}{\partial\alpha}(\alpha^2 N_1 + (1-\alpha)^2 N_2) = 0$, we get

$$2\alpha N_1 - 2(1-\alpha)N_2 = 0$$ [3 A]

$$\alpha = \frac{N_2}{N_1+N_2}$$

Capacity

$$C = \frac{1}{2}\log\left(1 + \frac{P}{N_1} + \frac{P}{N_2}\right)$$ [3 A]

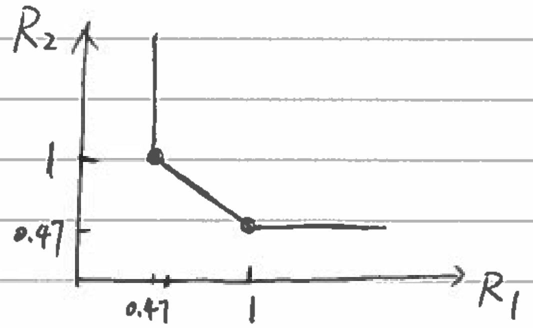
4.

a) Slepian – Wolf region

$$R_1 \geq H(X|Y)$$
$$R_2 \geq H(Y|X) \qquad [3B]$$
$$R_1 + R_2 \geq H(X,Y)$$

We have

$$H(X) = H(p) = 1$$
$$H(Y) = H(p * r) = 1 \text{ where } p*r = p(1-r) + r(1-p) = 0.5$$
$$H(X,Y) = H(X,Z)$$
$$= H(X) + H(Z)$$
$$= H(p) + H(r) = 1.47 \qquad [1E] \times 5 \text{ Each}$$
$$H(Y|X) = H(r) = H(Z) = 0.47$$
$$H(X|Y) = H(X,Y) - H(Y)$$
$$= H(p) + H(r) - H(p*r)$$
$$= 0.47$$



[1E]

b)

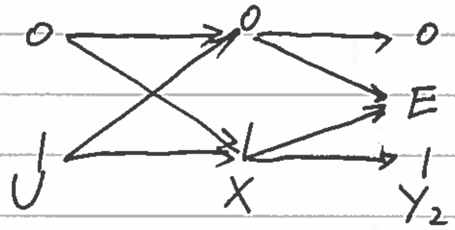
i) Since it is an erasure channel from X to Y_1,

$$C(X \to Y_1) = 1 - \alpha_1 \qquad [2A]$$

ii) It is easy to show that the link from X to Y_2 is an equivalent erasure channel with erasure probability $\alpha_1 + \alpha_2 - \alpha_1\alpha_2$. Hence [2A]

$$C(X \to Y_2) = (1-\alpha_1)(1-\alpha_2)$$
$$= 1 - \alpha_1 - \alpha_2 + \alpha_1\alpha_2 \qquad [2A]$$

iii) The equivalent channel from U to Y_2 is depicted as follows



It is easy to write down the transitional prob. [2T]

$U \backslash Y_2$	0	E	1
0	$(1-\beta)(1-\alpha)$	α	$\beta(1-\alpha)$
1	$\beta(1-\alpha)$	α	$(1-\beta)(1-\alpha)$

and the marginal distribution of Y_2

$$P_{Y_2}(y_2) = \left\{ \frac{1-\alpha}{2}, \alpha, \frac{1-\alpha}{2} \right\}$$

Hence, [1T]

$$R_2 = I(U; Y_2)$$
$$= H(Y_2) - H(Y_2|U)$$

Since $H(Y_2) = H\left(\frac{1-\alpha}{2}, \alpha, \frac{1-\alpha}{2}\right) = H(\alpha) + (1-\alpha)$

$$H(Y_2|U) = H((1-\beta)(1-\alpha), \alpha, \beta(1-\alpha))$$
$$= H(\alpha) + (1-\alpha)H(\beta)$$

We obtain

$$R_2 = (1-\alpha)(1-H(\beta)). \quad [2T]$$

Similarly,

$$R_1 = I(X; Y_1|U)$$
$$= H(Y_1|U) - H(Y_1|U,X)$$
$$= H(Y_1|U) - H(Y_1|X) \quad U \to X \to Y_1 \text{ is Markov chain}$$

We know

$$H(Y_1|X) = H(\alpha_1) \quad [1T]$$

Following the same procedure as that for $H(Y_2|U)$,

$$H(Y_1|U) = H(\alpha_1) + (1-\alpha_1)H(\beta)$$

Thus,

$$R_1 = (1-\alpha_1)H(\beta) \quad [2T]$$

Capacity region: [2T]