

440 Information theory.

The Solutions to Exam 2016

B—bookwork, E—new example, T—new theory

1.

- a) The possible outcomes of X and Y are given in the table below (O—Odd, E—Even):

X	1	2	3	4	5	6
Y	O	E	O	E	O	E

- i) Obviously,
 $H(X) = \log 6$, [2E]
 $H(Y) = 1$, [2E]
 because they are uniform.
- ii) $H(X|Y) = \frac{1}{2} H(X|Y=E) + \frac{1}{2} H(X|Y=O) = \frac{1}{2} \log 3 + \frac{1}{2} \log 3 = \log 3$ [2E]
 $H(Y|X) = \frac{1}{6} H(Y|X=1) + \frac{1}{6} H(Y|X=2) + \dots + \frac{1}{6} H(Y|X=6)$
 $= 0 + 0 + \dots + 0 = 0$ [2E]
 $H(X, Y) = H(X) + H(Y|X) = \log 6$
- iii) $I(X; Y) = H(X) - H(X|Y) = \log 6 - \log 3 = 1$ [2E]

b)

$$I(X; Y) = H(X) + H(Y) - H(X, Y)$$

$$= E \log \frac{P(X, Y)}{P(X)P(Y)} = D(P_{X,Y} \| P_X \otimes P_Y) \geq 0$$

[3B]

Equality holds iff $P(X, Y) = P(X)P(Y)$, i.e., X and Y are independent. [2B]

- c) Any $n - 1$ or fewer of these random variables are independent of each other.
 Thus, for $k \leq n - 1$,

$$I(X_{k+1}; X_k | X_1, X_2, \dots, X_{k-2}) = 0. \quad [3E]$$

However, given X_1, X_2, \dots, X_{n-2} , once we know either X_{n-1} or X_n , we know the other. Therefore, [2E]

$$I(X_{n-1}; X_n | X_1, X_2, \dots, X_{n-2})$$

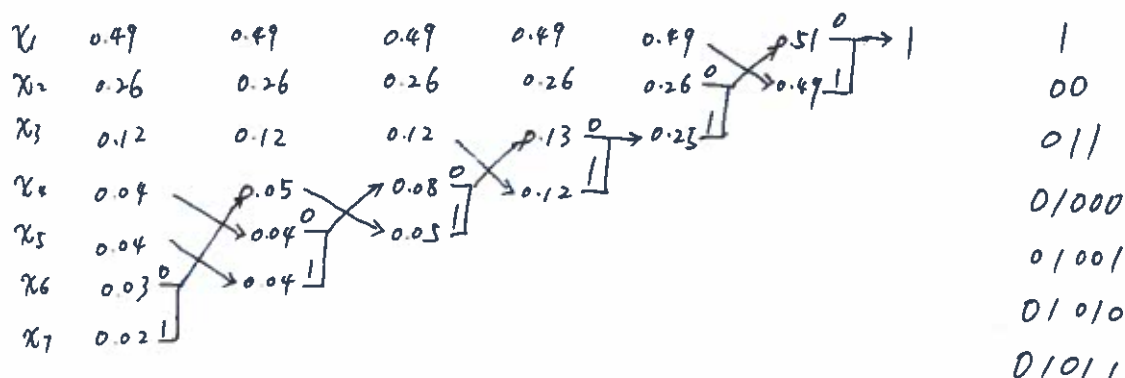
$$= H(X_n | X_1, X_2, \dots, X_{n-2}) - H(X_n | X_1, X_2, \dots, X_{n-1}) \quad [3E]$$

$$= 1 - 0 = 1 \text{ bit.} \quad [2E]$$

2.

a) i)

[5E]



ii)

$$L = \sum p(x_i) \ell(x_i) = 1 \times 0.49 + 2 \times 0.26 + 3 \times 0.12 + 5 \times 0.13 = 2.02$$

[3E]

b) Phrases: 1, 0, 10, 01, 100, 101, 00, 1000, 1010, 11

[2E]

Locations: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10

[2E]

There are 10 phrases, so 4 bits are needed to represent locations.

Encodings: (0000, 1); (0000, 0); (0001, 0); (0010, 1); (0011, 0)

(0011, 1); (0010, 0); (0101, 0); (0110, 0); (0001, 1)

[3E]

c) i)

The transition matrix is

$$P = \begin{pmatrix} 3/4 & 1/4 & 0 \\ 0 & 1/2 & 1/2 \\ 1/4 & 0 & 3/4 \end{pmatrix}$$

The stationary distribution is found from

$$\mu P = \mu$$

$$\Rightarrow \begin{cases} -\frac{1}{4}\mu_1 + \frac{1}{4}\mu_3 = 0 \\ \frac{1}{4}\mu_1 - \frac{1}{2}\mu_2 = 0 \\ \frac{1}{2}\mu_2 - \frac{1}{4}\mu_3 = 0 \end{cases}$$

Together with $\sum_i \mu_i = 1$ we get $\mu_1 = \frac{2}{5}$, $\mu_2 = \frac{1}{5}$, $\mu_3 = \frac{2}{5}$.

[3E]

The entropy rate is

$$\begin{aligned} H_{\infty}(U) &= \sum_i \mu_i H(S_i) = \frac{2}{5} h\left(\frac{1}{4}\right) + \frac{1}{5} h\left(\frac{1}{2}\right) + \frac{2}{5} h\left(\frac{1}{4}\right) \\ &= \frac{4}{5} \left(2 - \frac{3}{4} \log 3\right) + \frac{1}{5} = \frac{1}{9} - \frac{3}{4} \log 3 \approx 0.8490 \end{aligned} \quad [4E]$$

ii)

$$H\left(\frac{2}{5}, \frac{1}{5}, \frac{2}{5}\right) = -\frac{2}{5} \log \frac{2}{5} - \frac{1}{5} \log \frac{1}{5} - \frac{2}{5} \log \frac{2}{5} = \log 5 - \frac{4}{5} \approx 1.5219$$

That is, we gain in uncertainty if we take into consideration the memory of the source.

[3E]

3.

a)

- (1) definition of mutual info [1B]
- (2) Markov chain $w - x - y$, and data processing theorem [1B]
- (3) definition of mutual info [1B]
- (4) chain rule, translation-invariant [1B]
- (5) indep. bound, Fano's inequality [1B]
- (6) mutual info. \leq capacity [1B]
- (7) algebra [1B]
- (8) taking limit and $P_e^{(n)} \rightarrow 0$ [1B]

b)

i)

$$\begin{aligned}
 h(X) &= -\int_0^{\infty} f(x) \log f(x) dx = -\log e \int_0^{\infty} f(x) \ln f(x) dx \\
 &= -\log e \int_0^{\infty} \lambda e^{-\lambda x} (\ln \lambda - \lambda x) dx = -\log e (\ln \lambda - \int_0^{\infty} \lambda x e^{-\lambda x} d\lambda x) \\
 &= -\log e (\ln \lambda - 1) = -\log e \cdot \ln(\lambda / e) = \log(e / \lambda) \text{ bits}
 \end{aligned}$$

[2E]

[2E]

ii) The sum of two normal random variables is also normal, so applying the result derived in the class for the normal distribution,

$$h(f) = \frac{1}{2} \log 2\pi e(\sigma_1^2 + \sigma_2^2) \text{ bits}$$

[3E]

c) Recall the definition of $R(D)$:

$$R(D) = \min_{p(\hat{X}|X)} I(X; \hat{X}) \text{ such that } E[d(X, \hat{X})] \leq D$$

[2B]

We have

$$R(D) = \min h(X) - h(X|\hat{X}) \geq h(X) - \frac{1}{2} \log(2\pi e D)$$

[3T]

because

$$h(X|\hat{X}) = h(X - \hat{X}|\hat{X}) \leq h(X - \hat{X}) \leq \frac{1}{2} \log(2\pi e D)$$

[3T]

Therefore,

$$D \geq \frac{2^{2h(X)}}{2\pi e} 2^{-2R}$$

[2T]

4.

a)

Under no interference, each sender has power P and rate $C(P/N)$.

Each user independently sends a codeword from a Gaussian codebook. [2B]

Consider receiver 1 (receiver 2 is the same). Under very strong interference,

– Treats sender 1 as interference, and decode sender 2 at rate $C(a^2P/(P+N))$; [2B]

– Subtracting it from received signal, he sees a clean channel for sender 1 with capacity $C(P/N)$. [2B]

This is possible if $C(a^2P/(P+N)) > C(P/N)$, i.e., crosslink is better:

$$\frac{1}{2} \log \left(1 + \frac{P}{N} \right) < \frac{1}{2} \log \left(1 + \frac{a^2P}{P+N} \right) \quad [2B]$$

$$\frac{P}{N} < \frac{a^2P}{P+N}$$

$$a^2 > \frac{P+N}{N} = 1 + \frac{P}{N} \quad [2B]$$

b)

i)

[3E]

The joint distribution of X and Y is shown in following table

Z_1	Z_2	X	Y	probability
0	0	0	0	$(1-p_1)(1-p_2)$
0	1	1	-1	$(1-p_1)p_2$
1	0	1	1	$p_1(1-p_2)$
1	1	2	0	p_1p_2

and hence we can calculate

$$H(X) = H(p_1p_2, p_1 + p_2 - 2p_1p_2, (1-p_1)(1-p_2))$$

$$H(Y) = H(p_1p_2 + (1-p_1)(1-p_2), p_1 - p_1p_2, p_2 - p_1p_2)$$

and

$$H(X, Y) = H(Z_1, Z_2) = H(p_1) + H(p_2)$$

[4E]

Therefore,

$$H(X|Y) = H(p_1) + H(p_2) - H(p_1p_2 + (1-p_1)(1-p_2), p_1 - p_1p_2, p_2 - p_1p_2)$$

$$H(Y|X) = H(p_1) + H(p_2) - H(p_1p_2, p_1 + p_2 - 2p_1p_2, (1-p_1)(1-p_2))$$

Slepian-Wolf region is

$$R_X \geq H(X|Y)$$

$$R_Y \geq H(Y|X)$$

$$R_X + R_Y \geq H(X, Y) = H(p_1) + H(p_2)$$

[3E]

ii)

The Slepian Wolf region for (Z_1, Z_2) is

$$\begin{aligned} R_1 &\geq H(Z_1|Z_2) = H(p_1) \\ R_2 &\geq H(Z_2|Z_1) = H(p_2) \\ R_1 + R_2 &\geq H(Z_1, Z_2) = H(p_1) + H(p_2) \end{aligned}$$

which is a rectangular region.

[3E]

The minimum sum rate is the same in both cases, since if we knew both X and Y , we could find Z_1 and Z_2 and vice versa. However, the region in part (i) is usually pentagonal in shape, and is larger than the region in (ii).

[2E]