Information Theory 2018, Final Exam Solutions

Q1

a) 0.02

b) 0.95

c) 0.01

d) $P(\text{Albert took drugs} \mid \text{Test Positive})$

Write

$$P(A|T)\,P(T) = P(T|A)\,P(A)$$

$$\Rightarrow P(A|T) = \frac{P(T|A)\,P(A)}{P(T)}.$$

Now

$$P(T|A) = .95$$
$$P(A) = 0.02$$
$$P(T) = P(T|\neg A)\,P(\neg A) + P(T|A)\,P(A)$$
$$= 0.01 \times 0.98 + 0.95 \times 0.02$$
$$= \cancel{\epsilon \times (1-A) + \beta \times A}$$
$$= 0.0098 + 0.019$$
$$= 0.0288$$

So

$$P(A|T) = \frac{0.95 \times 0.02}{0.0288} = 0.6597$$

$$= 66\,\%.$$

Q 2

a) No. Require $P(X=x \text{ AND } Y=y)$
$= P(X=x)P(Y=y)$

b) $P(X=1) = P(X=1 \mid Y=1)\,P(Y=1)$
$+ P(X=1 \mid Y=0)\,P(Y=0)$

$< 0.2\,P(Y=1) + 0.6\,P(Y=0)$

$= \cancel{0.24}$

$\cancel{P(Y=1)\,[0.2}$

Do not know $P(Y=1)$. So all we know is $P(X=1) < 0.6$.
Thus $P(X=0) > 0.4$

c) $\cancel{P(X=1 \mid Y=0) = \frac{P(Y=0 \mid X=1)\,P(X=1)}{P(Y=0)}}$

$\cancel{P(Y=0 \mid X=1) = \frac{P(X=1 \mid Y=0)\,P(Y=0)}{P(X=1)}}$

$\cancel{\le \frac{0.6 \cdot P(Y=0)}{0.21}}$

c) $P(Y=0 \mid X=0) = P(X=0, Y=0)\,P(X=0)$ (a)
$P(Y=1 \mid X=0) = P(X=0, Y=1)\,P(X=0)$ (b)

~~Divide both sides~~

(a) $\Rightarrow$ $\dfrac{P(Y=0 \mid X=0)}{P(X=0, Y=0)} = P(X=0)$

(b) $\Rightarrow$ $\dfrac{P(Y=1 \mid X=0)}{P(X=0, Y=1)} = P(X=0)$

~~Thus~~ $\cancel{\dfrac{P(Y=1 \mid X=0)}{P(X=0, Y=0)} = \dfrac{P(Y=1 \mid X=0)}{P(X=1, Y=0)}}$

So $$\frac{P(Y=0 \mid X=0)}{P(X=0, Y=0)} = \frac{P(Y=1 \mid X=0)}{P(X=0, Y=1)}$$

$$\Rightarrow \frac{P(X=0, Y=1)}{P(X=0, Y=0)} = \frac{P(Y=1 \mid X=0)}{P(Y=0 \mid X=0)}$$

flip both sides & all are done.

5

$\cancel{Q4}$ 3). a). $H(X) = \log n = H(Y) = \cancel{H(Z)}$.
$H(Z) = \log_2 2n = H(Y) + 1$

b) Use decomposability:

$$H(\underline{p}) = H_2(p_1) + (1-p_1)\, H\left(\frac{p_2}{1-p_1}, \dots \frac{p_4}{1-p_1}\right)$$

$$= H_2(p_1) + (1-p_1)\left[H_2\left(\frac{p_2}{1-p_1}\right) + \left(1-\frac{p_2}{1-p_1}\right) H_2\left(\frac{\frac{p_3}{1-p_1}}{1-\frac{p_2}{1-p_1}}\right)\right]$$

~~c)~~

d)
1) No, Does not symmetric
2) Yes, because then $p=q \Rightarrow q=p$
3) No, just mean $p=q$
4) No, just mean $p=q$. Could be dep.
5) No. Same as 3
6) No, $I(X;Y)$ requires $X \perp Y$. See 4
7) Yes.
8) Yes $I(X;X) = H(X)$. Deterministic function does not change result.

~~3 e)~~ ~~4 a)~~ 3.d) Jensen along concave when

$$\mathbb{E}(f(x)) \leq f(\mathbb{E}(x))$$

Thus $\mathbb{E}\log(x) \leq 1$.

3 e) ~~f)~~. No solar panels means this consumption is always non-negative the Markov's $\neq$

$$P(X \geq \lambda) \leq \frac{E(X)}{\lambda}.$$

$\lambda = 2E(X)$ Thus probability is $\leq \frac{1}{2}$.

3 f ~~g)~~ Use Chebyshev

$$P(|X - E(X)| \geq \lambda\sqrt{V(X)}) \leq \frac{1}{\lambda^2}$$

$V(X) = 10 \Rightarrow$ Need $\lambda\sqrt{10} = 10$ so $\lambda = \sqrt{10}$

Thus $p \leq \frac{1}{10}$

6

Q4. a) $p(X) = (\frac{1}{8}, \frac{7}{8})$
$p(Y) = (3/4, 1/4)$
$p(X|Y=1) = (0, 1)$
$p(X|Y=2) = (\frac{1}{2}, \frac{1}{2})$

b) $H(X) = H_2(\frac{1}{8}) = -\frac{1}{8}\log_2\frac{1}{8} - \frac{7}{8}\log_2\frac{7}{8}$

$= \frac{1}{8}\cdot 3 - \frac{7}{8}\frac{\log_{10}(7/8)}{\log_{10}(2)}$

$= \frac{1}{8}\cdot 3 + \frac{7}{8}\cdot 0.1926$

$= .375 + .169$

$= 0.544$ ~~bits~~

$H(X|Y=1) = H_2(0) = 0$
$H(X|Y=2) = H_2(\frac{1}{2}) = 1 \geq 0.544$

c) $H(X|Y) = \sum_{y\in\{1,2\}} p(y)H(X|Y=y) = 0 + \frac{1}{4}\cdot 1 = \frac{1}{4} < H(X)$

Q5 a) {aaa, aab, aba, abb, baa, bab, bba, bbb}

b) abb, bba, or bab !
$T_{N\beta} = \{ \underline{x} \in A^3 \mid | -\frac{1}{3}\log_2 p(\underline{x}) - H(X) | < 0.2 \}$

$p(abb) = p(bba) = p(bab) = (\frac{2}{3})^2 \cdot \frac{1}{3} = \frac{4}{27}$

$-\frac{1}{3}\log_2(\frac{4}{27}) = 0.918$
$H(X) = H_2(\frac{1}{3}) = 0.917$ ✓.

c) Least probable seq is aaa w/ prob $\frac{1}{27}$
So $S_{\frac{1}{25}} = A \setminus \{a,a,a\}$
Next least prob is aab, aba, baa w/ prob $(\frac{1}{3})^2 \cdot \frac{2}{3} = \frac{2}{27}$. Need 2 of these to make sum $> \frac{1}{10}$

d) $|X^{1000}| = 2^{1000}$. $|T_{N\beta}| = 2^{NH(x)+\beta} = 2^{1000 \times 0.9182}$

$\approx 2^{918.2}$

Ratio is $\approx 2^{-82}$ — that is $\approx$ the fraction of sequences that are typical!

e) By the asymptotic equipartition property the probability is close to 1. (!)

Q6. a) $4 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + \frac{1}{8} \cdot 2 + \frac{1}{8} \cdot 1 = 2\frac{7}{8}$

b) No, 0 is a prefix of 0000

c) No, 0000 can be decoded as a

d) $H(X) = \frac{1}{2}\log_2 2 + \frac{1}{4}\log_2 4 + \frac{2}{8}\log_2 8$

$= \frac{1}{2} + \frac{1}{2} + \frac{6}{8}$

$= 1\frac{3}{4} = 1.75$

e) Use Kraft inequality: Require $\sum_i 2^{-l_i} \leq 1$

We have $\frac{1}{2} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} > 1$, so no.

f) Shannon code word lengths are $l_i = \lceil \log_2 \frac{1}{p_i} \rceil$

So $l_1 = 1$ eg. 0

$l_2 = 2$ 10

$l_3 = 3$ 110

$l_4 = 4$ 111

g)

a	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$ — 0	1
b	1/4	$\frac{1}{4}$ — 0	$\frac{1}{2}$ / 1	
c	1/8 — 0	$\frac{1}{4}$ / 1		
d	1/8 / 1			

$\Rightarrow \{0, 10, 110, 111\}$ Same!

h) Extended ensemble. Non-symbol codes

7) a) lengths =, not prefix free
b) lengths =
c) Yes.
d) Code extended ensemble
e) Looking for key orders

10

8

a) If (X, Y, Z) is a Markov chain, then $X \perp Z \mid Y$ and thus

$$p(X, Z \mid Y) = p(X \mid Y)\, p(Z \mid Y)$$

Thus $p(Z, Y, X) = p(Z)\, p(Y \mid Z)\, p(X \mid Y, Z)$

$$= p(Z)\, p(Y \mid Z)\, p(X \mid Y)$$

since $X \perp Z \mid Y$. Thus (Z, Y, X) is a Markov chain

b) The data processing inequality implies that

$$I(X; Y) \geq I(X; Z)$$

and using a), we also have

$$I(Z; Y) \geq I(Z; X)$$

But $I(Z; X) = I(X; Z)$ and so

$$I(X; Z) \leq I(Z; Y)$$

and $I(X; Z) \leq I(X; Y)$.

The result follows.

c) i) $p(x, y, z) = p(z \mid y)\, p(y \mid x)\, p(x)$ since $X \perp Z \mid Y$

ii) Yes by i)

iii) Yes by the data processing $\neq$

iv) ~~Part b) showed $I(X,Z) \leq \min(I(X,Y), I(Y,Z))$~~

~~The capacity of C^{XZ}~~

$= \cancel{\min_p I(X,Z) \leq \min_p \min(I(X,Y), I(Y,Z))}$ (*)

$\leq \min(C^{XY}, C^{YZ})$ ~~(The argmin for (*) might not be the argmin for capacity.)~~

(iv) $$C^{XZ} = \min_p I(X;Z)$$
$$\le \min_p \min\left(I(X;Y), I(Y;Z)\right)$$

The minimization ~~and~~ over p will either minimize $I(X;Y)$, $I(Y;Z)$ or both.

If it minimizes $I(X;Y)$, then $I_{p_Y}(Y;Z)$ must be greater that $C^{XY} = \min_p I(X;Y)$.

Here p_Y is the distribution of Y induced by X. This is true because otherwise the $\min_p$ would have minimized $I(Y;Z)$.

In this case we get
$$C^{XZ} \le \min\left(C^{XY}, I_{p_Y}(Y;Z)\right)$$

The other case (similar argument) implies
$$C^{XZ} \le \min\left(I_p(X;Y), C^{YZ}\right).$$

~~If $\min_p$ minimises both~~

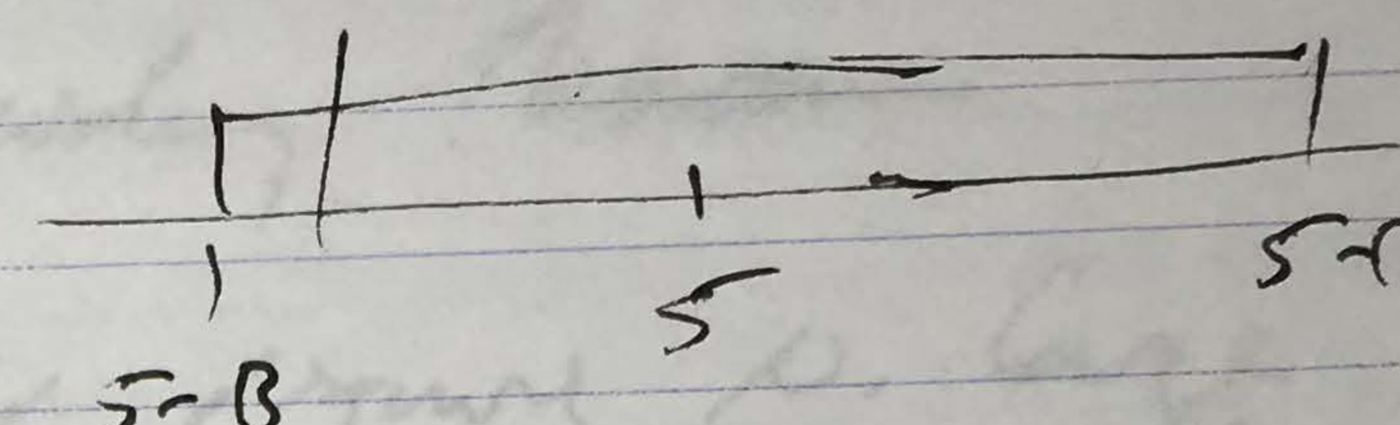
Then
$$C^{XZ} \le \min\left[\min\left(C^{XY}, I_{p_Y}(Y;Z)\right), \min\left(I_p(X;Y), C^{YZ}\right)\right]$$
$$= \min\left(C^{XY}, C^{YZ}\right).$$

Intuition is simple serial connection – thinnest-pipe limits capacity of the whole.

Q9 a) By symmetry there is a binary symmetric channel $Q=\begin{bmatrix}1-f & f \\ f & 1-f\end{bmatrix}$

where $f = p(\hat{X}=1 \mid \hat{X}=\cancel{0}-1) = p(\hat{X}=-1 \mid X=1)$

$$= \int_{5-B}^{0} \frac{1}{2B}\,dx \quad \text{when } B>5$$



$$= \frac{B-5}{2B} \quad (B>5)$$

$$0 \quad 0 \quad (B \leq 5)$$

b) ~~then B~~ when $B \leq 5$, $f=0$ and $C^{XY}=1$.

Now let us just determine the capacity of the binary symmetric channel.

~~By~~ We require $\max\limits_{p} I(X;Y)$

where p is the distribution of X. Since Q^{XY} is symmetric, the optimal distribution is uniform! We thus just need to compute $I(X;Y)$ when $p_X = (\underset{p_0}{\tfrac{1}{2}}, \underset{p_1}{\tfrac{1}{2}})$.

Use $I(X;Y) = H(Y) - H(Y|X)$

$$H(Y) = H_2(f p_0 + (1-f) p_1)$$
$$= H_2\left(f\cdot\tfrac{1}{2} + (1-f)\cdot\tfrac{1}{2}\right) = H_2\left(\tfrac{1}{2}\right) = 1.$$

$$H(Y|X) = \sum_{x} p(X=x)\, H(f) = H_2(f).$$

Thus $C^{XY} = H(Y) - H(Y|X)$
$= 1 - H_2(f)$

So when $B > S$, $C^{XY} = 1 - H_2\left(\frac{B-S}{2B}\right)$

---

c) Yes, channel coding theorem

d) No CCT would require v. large block length to get this close to the capacity.
Looking for which look jointly typically,
AEP all extended channels look like noisy ~~tyf~~ when one uses large enough block sizes.

e). There will be $\frac{1}{1024}$ (2^{-10}) !