IMPERIAL COLLEGE LONDON

EE4-40
EE9-CS7-26
EE9-SO20

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
EXAMINATIONS 2019

MSc and EEE/EIE PART IV: MEng and ACGI

Corrected copy

INFORMATION THEORY

Correction 10:55
2c

Thursday, 9 May 10:00 am

Time allowed: 3:00 hours

There are FOUR questions on this paper.

Answer ALL questions.

All questions carry equal marks

Any special instructions for invigilators and information for candidates are on page 1.

Examiners responsible First Marker(s) : C. Ling
Second Marker(s) : D. Gunduz

Information for students

Notation:

(a) Random variables are shown in Tahoma font. X, $\mathbf{x}$, $\mathbf{X}$ denote a random scalar, vector and matrix respectively.

(b) The size of a set A is denoted by $|A|$.

(c) By default, the logarithm is to the base 2.

(d) $\oplus$ denotes the exclusive-or operation, or modulo-2 addition.

(e) "i.i.d." means "independent identically distributed".

(f) $H(\cdot)$ is the entropy function.

(g) $C(x) = \frac{1}{2}\log_2(1+x)$ is the capacity function for the Gaussian channel in bits/channel use.

The Questions

1. Basics of information theory.

a) Suppose X_1 and X_2 are i.i.d. Bernoulli random variables taking values of 0 and 1 with equal probabilities ($p = 0.5$). Let $Y = \max(X_1, X_2)$. Compute the following entropy or mutual information:

i) $H(Y)$

ii) $I(X_1 ; Y)$

iii) $I(X_{1:2} ; Y)$

[9]

b) For $\mathbf{p} = [\frac{1}{3}, \frac{2}{3}]$ and $\mathbf{q} = [\frac{1}{2}, \frac{1}{2}]$, compute the relative entropy terms $D(\mathbf{p}||\mathbf{q})$ and $D(\mathbf{q}||\mathbf{p})$.

[6]

c) Data processing theorem for relative entropy. Suppose after going through a channel with transitional probabilities $W(y \mid x)$, distribution $\mathbf{p}$ becomes $\mathbf{p}'$, while $\mathbf{q}$ becomes $\mathbf{q}'$. Show that

$$D(\mathbf{p}'||\mathbf{q}') \leq D(\mathbf{p}||\mathbf{q})$$

Hint: use the so-called log-sum inequality (which can be derived from Jensen's inequality)

$$\sum_{i=1}^{n} a_i \log \frac{a_i}{b_i} \geq \left(\sum_{i=1}^{n} a_i \right) \log \frac{\sum_{i=1}^{n} a_i}{\sum_{i=1}^{n} b_i}$$

where a_i and b_i's are nonnegative numbers.

[10]

2. Fano's inequality.

Consider the Markov chain shown in Fig. 2.1, where X and Y are discrete random variables, and $\hat{X}$ is the estimate of X.

$$X \longrightarrow Y \longrightarrow \hat{X}$$

Fig. 2.1. Markov chain arising in Fano's inequality.

a) Define a random variable $e = (\hat{X} \neq X) \in \{0,1\}$. Justify each step of the following derivations.

$$H(e,X \mid Y) \overset{(1)}{=} H(X \mid Y) + H(e \mid X,Y) \overset{(2)}{=} H(e \mid Y) + H(X \mid e,Y)$$

$$\overset{(3)}{\Rightarrow} H(X \mid Y) + 0 \leq H(e) + H(X \mid e,Y)$$

$$\overset{(4)}{=} H(e) + H(X \mid Y, e=0)(1-p_e) + H(X \mid Y, e=1)p_e$$

$$\overset{(5)}{\leq} H(p_e) + 0 \times (1-p_e) + \log(|\mathsf{X}|-1)p_e$$

$$\overset{(6)}{\Rightarrow} p_e \geq \frac{(H(X \mid Y) - H(p_e))}{\log(|\mathsf{X}|-1)} \overset{(7)}{\geq} \frac{(H(X \mid Y) - 1)}{\log(|\mathsf{X}|-1)}$$

[8]

b) Given the following joint distribution

X \ Y	a	b
1	1/3	1/12
2	1/12	1/3
3	1/12	1/12

Find the minimum error probability corresponding to the optimum estimator of X given Y, and compare it with Fano's inequality for this problem.

[10]

c) Suppose X is a finite-entropy random variable defined on an infinite set X, e.g., the integers $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$. How would you modify Fano's inequality in this case?

[7]

3. Gaussian sources and channels.

a) Justify each step in the following derivation of the entropy of multivariate Gaussian sources.

Given mean, **m**, and symmetric positive definite covariance matrix **K**, we have the probability density function

$$X_{1:n} \sim \mathrm{N}(\mathbf{m},\mathbf{K}) \quad \Leftrightarrow \quad f(\mathbf{x}) \overset{(1)}{=} |2\pi\mathbf{K}|^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{x}-\mathbf{m})^T\mathbf{K}^{-1}(\mathbf{x}-\mathbf{m})\right)$$

Then,

$$h(f) \overset{(2)}{=} -(\log e)\int f(\mathbf{x})\times\left(-\frac{1}{2}(\mathbf{x}-\mathbf{m})^T\mathbf{K}^{-1}(\mathbf{x}-\mathbf{m}) - \tfrac{1}{2}\ln|2\pi\mathbf{K}|\right)d\mathbf{x}$$

$$\overset{(3)}{=} \tfrac{1}{2}\log(e)\times\left(\ln|2\pi\mathbf{K}| + E\left((\mathbf{x}-\mathbf{m})^T\mathbf{K}^{-1}(\mathbf{x}-\mathbf{m})\right)\right)$$

$$\overset{(4)}{=} \tfrac{1}{2}\log(e)\times\left(\ln|2\pi\mathbf{K}| + E\,\mathrm{tr}\left((\mathbf{x}-\mathbf{m})(\mathbf{x}-\mathbf{m})^T\mathbf{K}^{-1}\right)\right)$$

$$\overset{(5)}{=} \tfrac{1}{2}\log(e)\times\left(\ln|2\pi\mathbf{K}| + \mathrm{tr}\left(E(\mathbf{x}-\mathbf{m})(\mathbf{x}-\mathbf{m})^T\mathbf{K}^{-1}\right)\right)$$

$$\overset{(6)}{=} \tfrac{1}{2}\log(e)\times\left(\ln|2\pi\mathbf{K}| + \mathrm{tr}\left(\mathbf{K}\mathbf{K}^{-1}\right)\right)$$

$$\overset{(7)}{=} \tfrac{1}{2}\log(e)\times\left(\ln|2\pi\mathbf{K}| + n\right) \overset{(8)}{=} \tfrac{1}{2}\log\left(e^n\right) + \tfrac{1}{2}\log\left(|2\pi\mathbf{K}|\right)$$

$$\overset{(9)}{=} \tfrac{1}{2}\log\left(|2\pi e\mathbf{K}|\right) \overset{(10)}{=} \tfrac{1}{2}\log\left((2\pi e)^n|\mathbf{K}|\right) \text{ bits}$$

[10]

b) Consider an additive-noise fading channel

$$Y = VX + Z$$

where Z is additive noise, V is a random variable representing the fading coefficient, and Z and V are independent of each other and of X. Show that the knowledge of fading coefficient V improves the capacity:

$$I(X;Y\mid V) \geq I(X;Y).$$

[8]

c) Consider the additive Gaussian noise channel

$$Y_i = X_i + Z_i$$

where Z_i's are independent with mean μ_i and variance N. X_i's have power constraint P. Find the capacity if

i) $\mu_i = 0$, for all i.

ii) μ_i is unknown to the transmitter or receiver, but i.i.d. Gaussian with mean 0 and variance N_1 for all i.

[7]

4. Network information theory.

a) Multi-access channel.

i) With the help of a diagram, explain the term multi-access channel. Give an example.

ii) Write down the capacity region of the two-user bandlimited multi-access channel with additive white Gaussian noise, where the users have equal powers P, bandwidth W, and the noise power is N. Draw the region and explain how to achieve the two corner points.

iii) Find the capacity region for this two-user multiple access channel with infinite bandwidth W. Show that both senders can send at their individual capacities, i.e., infinite bandwidth eliminates interference.

[15]

b) Parallel transmission. Consider a scenario where the transmitted signal X of power P is received by two antennas:

$$Y_1 = \alpha X + Z_1$$
$$Y_2 = (1-\alpha)X + Z_2$$

where $0<\alpha<1$, Z_1 and Z_2 are independent Gaussian noises of zero mean and power N.

i) Assuming that both signals Y_1 and Y_2 are available at a common decoder $\mathbf{Y}=(Y_1; Y_2)$; what is the capacity of the channel from the sender to the common decoder?

ii) If instead the two receivers Y_1 and Y_2 each independently decode their signals, this becomes a broadcast channel. Let R_1 be the rate to base station 1 and R_2 be the rate to base station 2. Find the capacity region of this channel.

[10]