

COM5120 Communication theory

Homework #4 2021

Reference solution

1. Entropy, Information and Capacity

$$(1) X \sim \left(\frac{1}{4}, \quad \frac{1}{4}, \quad \frac{1}{4}, \quad \frac{1}{4} \right)$$

$$H(X) = \left(-\frac{1}{4} \times \log \frac{1}{4} \right) * 4 = 2 \text{ (bits)}$$

$$Y \sim \left(\frac{1}{2}, \quad \frac{1}{4}, \quad \frac{1}{8}, \quad \frac{1}{8} \right)$$

$$H(Y) = \left(-\frac{1}{2} \times \log \frac{1}{2} \right) + \left(-\frac{1}{4} \times \log \frac{1}{4} \right) + \left(-\frac{1}{8} \times \log \frac{1}{8} \right) + \left(-\frac{1}{8} \times \log \frac{1}{8} \right)$$

$$= \frac{1}{2} + \frac{1}{2} + \frac{3}{8} + \frac{3}{8} = \frac{7}{4} \text{ (bits)}$$

$$(2) H(X, Y) = 1 \times \left(-\frac{1}{4} \times \log \frac{1}{4} \right) + 2 \times \left(-\frac{1}{8} \times \log \frac{1}{8} \right) + 6 \times \left(-\frac{1}{16} \times \log \frac{1}{16} \right)$$

$$+ 4 \times \left(-\frac{1}{32} \times \log \frac{1}{32} \right)$$

$$= 1 \times \frac{2}{4} + 2 \times \frac{3}{8} + 6 \times \frac{4}{16} + 4 \times \frac{5}{32}$$

$$= \frac{27}{8} \text{ (bits)}$$

$$(3) H(X|Y) = H(X, Y) - H(Y) = \frac{27}{8} - \frac{7}{4} = \frac{13}{8} \text{ (bits)}$$

$$H(Y|X) = H(X, Y) - H(X) = \frac{27}{8} - 2 = \frac{11}{8} \text{ (bits)}$$

$$(4) I(X, Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$$

$$= 2 - \frac{13}{8} = \frac{7}{4} - \frac{11}{8} = \frac{3}{8} \text{ (bits)}$$

$$(5) \mathcal{C} = \max(I(X, Y)) = \frac{3}{8} \text{ (bits)}$$

2. Differential entropy of the continuous random variable

$$\begin{aligned}
 (1) \quad H(X) &= - \int_0^{\infty} \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \ln \left(\frac{1}{\lambda} e^{-\frac{x}{\lambda}} \right) dx \\
 &= - \ln \frac{1}{\lambda} \int_0^{\infty} \frac{1}{\lambda} e^{-\frac{x}{\lambda}} dx + \int_0^{\infty} \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \frac{x}{\lambda} dx \\
 &= \ln \lambda + \frac{1}{\lambda} \int_0^{\infty} \frac{1}{\lambda} e^{-\frac{x}{\lambda}} x dx \\
 &= \ln \lambda + \frac{1}{\lambda} \lambda = 1 + \ln \lambda
 \end{aligned}$$

$$\begin{aligned}
 (2) \quad H(X) &= - \int_{-\infty}^{\infty} \frac{1}{2\lambda} e^{-\frac{|x|}{\lambda}} \ln \left(\frac{1}{2\lambda} e^{-\frac{|x|}{\lambda}} \right) dx \\
 &= - \ln \frac{1}{2\lambda} \int_{-\infty}^{\infty} \frac{1}{2\lambda} e^{-\frac{|x|}{\lambda}} dx + \frac{1}{\lambda} \int_{-\infty}^{\infty} |x| \frac{1}{2\lambda} e^{-\frac{|x|}{\lambda}} dx \\
 &= \ln(2\lambda) + \frac{1}{\lambda} \left[\int_{-\infty}^0 -x \frac{1}{2\lambda} e^{\frac{x}{\lambda}} dx + \int_0^{\infty} x \frac{1}{2\lambda} e^{-\frac{x}{\lambda}} dx \right] \\
 &= \ln(2\lambda) + \frac{1}{2\lambda} \lambda + \frac{1}{2\lambda} \lambda = 1 + \ln(2\lambda)
 \end{aligned}$$

3. AWGN channel: $Y = X + N$

(1)

$$H(X|N) = - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x,n) \log_2 p(x|n) dx dn$$

Since X, N are independent, $p(x,n) = p(x)p(n)$, $p(x|n) = p(x)$. Hence,

$$\begin{aligned}
 H(X|N) &= - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x)p(n) \log_2 p(x) dx dn = - \int_{-\infty}^{\infty} p(n) dn \int_{-\infty}^{\infty} p(x) \log_2 p(x) dx \\
 &= H(X) = \frac{1}{2} \log_2 (2\pi e \sigma_x^2)
 \end{aligned}$$

(2)

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

Since Y is the sum of two independent, zero-mean Gaussian r.v.'s, it is also a zero-

mean Gaussian r.v. with variance $\sigma_x^2 + \sigma_n^2$. Hence, $H(Y) = \frac{1}{2} \log_2(2\pi e(\sigma_x^2 + \sigma_n^2))$.

$$\text{Also, since } Y = X + N, p(y|x) = p_n(y-x) = \frac{1}{\sqrt{2\pi\sigma_n^2}} e^{-\frac{(y-x)^2}{2\sigma_n^2}}$$

Hence,

$$\begin{aligned} H(Y|X) &= -\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x,y) \log_2 p(y|x) dx dy \\ &= -\int_{-\infty}^{\infty} p(x) \log_2 e \int_{-\infty}^{\infty} p(y|x) \ln \left(\frac{1}{\sqrt{2\pi\sigma_n^2}} \exp \left(-\frac{(y-x)^2}{2\sigma_n^2} \right) \right) dy dx \\ &= \int_{-\infty}^{\infty} p(x) \log_2 e \left[\int_{-\infty}^{\infty} p_n(y-x) \left(\ln(\sqrt{2\pi\sigma_n^2}) + \frac{(y-x)^2}{2\sigma_n^2} \right) dy \right] dx \\ &= \int_{-\infty}^{\infty} p(x) \log_2 e \left[\ln(\sqrt{2\pi\sigma_n^2}) + \frac{1}{2\sigma_n^2} \sigma_n^2 \right] dx \\ &= \left[\frac{1}{2} \log_2(2\pi\sigma_n^2) + \frac{1}{2} \log_2(e) \right] \int_{-\infty}^{\infty} p(x) dx \\ &= \frac{1}{2} \log_2(2\pi e \sigma_n^2) = H(N) \end{aligned}$$

Note that: $\int_{-\infty}^{\infty} p_n(y-x) dy = 1$, $\int_{-\infty}^{\infty} p_n(y-x)(y-x)^2 dy = E[N^2] = \sigma_n^2$

From $H(Y)$ and $H(Y|X)$:

$$I(X;Y) = H(Y) - H(Y|X) = \frac{1}{2} \log_2(2\pi e(\sigma_x^2 + \sigma_n^2)) - \frac{1}{2} \log_2(2\pi e \sigma_n^2) = \frac{1}{2} \log_2 \left(1 + \frac{\sigma_x^2}{\sigma_n^2} \right).$$

4. Entropy and information of joint random variables

$$(1) H(X, Y | Z) \geq H(X|Z)$$

$$H(X, Y | Z) = H(X|Z) + H(Y | X, Z) \geq H(X|Z).$$

with equality iff $H(Y | X, Z) = 0$

$$(2) I(X, Y; Z) \geq I(X; Z | Y)$$

$$I(X, Y; Z) = I(X; Z | Y) + I(Y; Z) \geq I(X; Z | Y)$$

with equality iff $I(Y; Z) = 0$

$$(3) H(X, Y, Z) - H(X | Y, Z) \geq H(Z | X, Y)$$

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

with equality iff $H(Y) + I(X; Z | Y) = 0$

$$(4) I(X; Z | Y) \geq I(Z; Y | X) - I(Z; Y) - I(X; Z)$$

$$\text{Since } I(Z; Y) = I(Z; Y | X) + I(X; Y; Z)$$

$$\text{, and } I(X; Z) = I(X; Z | Y) + I(X; Y; Z).$$

Hence, we can rewrite the inequality as

$$I(X; Z) - I(X; Y; Z)$$

$$\geq I(Z; Y) - I(X; Y; Z) - I(Z; Y) - I(X; Z), \text{ then we get}$$

$$I(X; Z) \geq 0.$$

Therefore, this is true in all cases.

And the equality holds iff $I(X; Z) = 0$.

5. Z channel

(1) (5%) $I(X, Y) = H(Y) - H(Y|X)$, and assume $p(0) = q$, $p(1) = 1 - q$.

$$\begin{aligned} H(Y|X) &= \sum_x P(x) H(Y|X=x) \\ &= q H(Y|X=0) + (1-q) H(Y|X=1) \\ &= q \times 0 + (1-q) H(Y|X=1) \end{aligned}$$

Where

$$\begin{aligned} H(Y|X=1) &= - \sum_y P(Y=y|X=1) \log_2 P(Y=y|X=1) \\ &= -P(Y=0|X=1) \log_2 P(Y=0|X=1) - P(Y=1|X=1) \log_2 P(Y=1|X=1) \\ &= -(1-\varepsilon) \log_2 (1-\varepsilon) - \varepsilon \log_2 \varepsilon \\ &= H(\varepsilon) \end{aligned}$$

$$\therefore H(Y|X) = (1-q) H(\varepsilon)$$

The probability mass function of the output symbols is:

$$\begin{aligned} P(Y=0) &= P(X=0)P(Y=0|X=0) + P(X=1)P(Y=0|X=1) \\ &= q \times 1 + (1-q) \times (1-\varepsilon) \\ &= 1 - \varepsilon + q\varepsilon \\ P(Y=1) &= P(X=0)P(Y=1|X=0) + P(X=1)P(Y=1|X=1) \\ &= q \times 0 + (1-q) \times \varepsilon \\ &= \varepsilon - q\varepsilon \end{aligned}$$

$$\therefore H(Y) = H(\varepsilon - q\varepsilon)$$

Hence, the capacity C :

$$C = \max_q I(X, Y) = \max_q [H(Y) - H(Y|X)] = \max_q [H(\varepsilon - q\varepsilon) - (1-q)H(\varepsilon)]$$

$$\Rightarrow \frac{\partial C}{\partial q} = -\varepsilon \log_2 (1 - \varepsilon + q\varepsilon) + \varepsilon \log_2 (\varepsilon - q\varepsilon) + H(\varepsilon) = 0$$

$$\Rightarrow H(\varepsilon) = -\varepsilon \log_2 \frac{\varepsilon - q\varepsilon}{1 - \varepsilon + q\varepsilon}$$

$$\Rightarrow q = \frac{\varepsilon + 2^{-\frac{H(\varepsilon)}{\varepsilon}} (\varepsilon - 1)}{\varepsilon \times \left(1 + 2^{-\frac{H(\varepsilon)}{\varepsilon}}\right)}$$

Plug in and derive the capacity:

$$C = H\left(\frac{2^{-\frac{H(\varepsilon)}{\varepsilon}}}{1 + 2^{-\frac{H(\varepsilon)}{\varepsilon}}}\right) - \frac{H(\varepsilon) \times 2^{-\frac{H(\varepsilon)}{\varepsilon}}}{\varepsilon \times \left(1 + 2^{-\frac{H(\varepsilon)}{\varepsilon}}\right)} = C(\varepsilon)$$

(2) (10%) If $\varepsilon = 0$, by L'Hospital's rule: $\lim_{\varepsilon \rightarrow 0} \frac{H(\varepsilon)}{\varepsilon} = \infty$, $\lim_{\varepsilon \rightarrow 0} \frac{H(\varepsilon)}{\varepsilon} 2^{-\frac{H(\varepsilon)}{\varepsilon}} = 0$.

and the channel capacity: $\lim_{\varepsilon \rightarrow 0} C(\varepsilon) = H(0) = 0$.

If $\varepsilon = 1$, then $C = H\left(\frac{2^{-H(1)}}{1+2^{-H(1)}}\right) - \frac{H(1) \times 2^{-H(1)}}{1+2^{-H(1)}} = H\left(\frac{1}{2}\right) = 1$.

and $q = \frac{1}{(1+2^{-H(1)})} = \frac{1}{2}$.

Hence, the input distribution that achieves capacity is $p(0) = \frac{1}{2}, p(1) = \frac{1}{2}$.

If $\varepsilon = 0.57$, then $C = H\left(\frac{2^{-\frac{H(0.57)}{0.57}}}{1+2^{-\frac{H(0.57)}{0.57}}}\right) - \frac{H(0.57) \times 2^{-\frac{H(0.57)}{0.57}}}{0.57 \times \left(1+2^{-\frac{H(0.57)}{0.57}}\right)}, H(0.57) = 0.9858$.

$\therefore C = H\left(\frac{0.3016}{1.3016}\right) - 0.4007 = H(0.2317) - 0.4007 = 0.781 - 0.4007 = 0.3803$.

and $q = \frac{0.57 + 2^{-\frac{H(0.57)}{0.57}} (0.57 - 1)}{0.57 \times \left(1+2^{-\frac{H(0.57)}{0.57}}\right)} = 0.5935$. Hence, the input distribution that

achieves capacity is $p(0) = 0.5935, p(1) = 0.4065$.

(3) (5%) The conditional probability:

$$\begin{aligned} P(Y=0|X=1) &= (1-\varepsilon) + \varepsilon(1-\varepsilon) + \varepsilon^2(1-\varepsilon) + \dots \\ &= (1-\varepsilon)(1+\varepsilon+\varepsilon^2+\dots) \\ &= (1-\varepsilon) \frac{1-\varepsilon^n}{1-\varepsilon} \\ &= 1-\varepsilon^n \end{aligned}$$

The resulting system is equivalent to a Z channel with $\varepsilon_1 = \varepsilon^n$.

(4) (5%)

Case 1: $0 \leq \varepsilon < 1$

As $n \rightarrow \infty$, $\varepsilon_1 = \varepsilon^n \rightarrow 0$. Thus, from (2) the channel capacity $C \rightarrow 0$.

Case 2: $\varepsilon = 1$

As $n \rightarrow \infty$, $\varepsilon_1 = \varepsilon^n \rightarrow 1$. Thus, from (2) the channel capacity $C \rightarrow 1$.