Sample problems for the 9th lecture (May 10)

1. (a) Suppose that the continuous random variable X has probability density function (pdf) $f_X(x)$. Express the probability mass function of $q_n(X)$ in terms of $f_X(\cdot)$, where

$$q_n(x) := m\Delta \quad \text{for } m\Delta \leq x < (m+1)\Delta$$

is the uniform quantization on $x \in \mathbb{R}$.

(b) Express the entropy of $q_n(X)$ in terms of $f_X(\cdot)$.

(c) Let $0 \leq f_X(x) \leq A$ be a bounded function for $x \in \mathbb{R}$. Show that the entropy of $q_n(X)$ is lower bounded by

$$\log_2 \frac{1}{\Delta A}.$$

Hint: $\displaystyle\int_{m\Delta}^{(m+1)\Delta} f_X(x)dx \leq \Delta A$.

(d) Prove that the entropy of X is infinity.

Hint: $H(f(X)) \leq H(X)$ for any function f.

Solution.

(a) $\displaystyle\Pr[q_n(X) = m\Delta] = \int_{m\Delta}^{(m+1)\Delta} f_X(x)dx$

(b)

$$\begin{aligned}
H(q_n(X)) &= \sum_{m=-\infty}^{\infty} \Pr[q_n(X) = m\Delta] \log_2 \frac{1}{\Pr[q_n(X) = m\Delta]} \\
&= \sum_{m=-\infty}^{\infty} \left(\int_{m\Delta}^{(m+1)\Delta} f_X(x)dx \right) \log_2 \frac{1}{\left(\int_{m\Delta}^{(m+1)\Delta} f_X(x)dx \right)}
\end{aligned}$$

(c) From (b),

$$\begin{aligned}
H(q_n(X)) &= \sum_{m=-\infty}^{\infty} \left(\int_{m\Delta}^{(m+1)\Delta} f_X(x)dx \right) \log_2 \frac{1}{\left(\int_{m\Delta}^{(m+1)\Delta} f_X(x)dx \right)} \\
&\geq \sum_{m=-\infty}^{\infty} \left(\int_{m\Delta}^{(m+1)\Delta} f_X(x)dx \right) \log_2 \frac{1}{\Delta A} \\
&= \log_2 \frac{1}{\Delta A} \sum_{m=-\infty}^{\infty} \left(\int_{m\Delta}^{(m+1)\Delta} f_X(x)dx \right) \\
&= \log_2 \frac{1}{\Delta A} \int_{-\infty}^{\infty} f_X(x)dx \\
&= \log_2 \frac{1}{\Delta A}.
\end{aligned}$$

(d) From (c), $H(X) \geq H(q_n(X)) \geq \log_2 \frac{1}{\Delta A}$; hence,

$$H(X) \geq \lim_{\Delta \downarrow 0} \log_2 \frac{1}{\Delta A} = \infty.$$

2. Determine the differential entropy (in nats) of random variable X for each of the following cases.

   (a) X is exponential with parameter λ, i.e., $f_X(x) = \lambda e^{-\lambda x}, x \geq 0$.

   (b) X is Laplacian with parameter λ and mean zero, i.e., $f_X(x) = \frac{1}{2\lambda} e^{-\frac{|x|}{\lambda}}, x \in \mathbb{R}$.

   (c) $X = aX_1 + bX_2$, where a and b are non-zero constants and X_1 and X_2 are independent Gaussian random variables such that $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$.
   Hint: The differential entropy of Gaussian random variable with variance σ^2 is $h(X) = \frac{1}{2} \ln (2\pi e \sigma^2)$.

Solution:

(a) Exponential distribution: $f_X(x) = \lambda e^{-\lambda x}$ if $x \geq 0$.

$$\begin{aligned} h(X) &= -\int_0^\infty f_X(x) \ln f_X(x) dx \\ &= -\int_0^\infty \lambda e^{-\lambda x} (\ln \lambda - \lambda x)\, dx \\ &= \ln\left(\frac{e}{\lambda}\right) \\ &= 1 - \ln(\lambda) \text{ nats} \end{aligned}$$

where we have used the fact that $E[X] = 1/\lambda$.

(b) Laplace distribution: $f_X(x) = (1/2\lambda)e^{-\frac{|x|}{\lambda}}$, $x \in \mathbb{R}$.

$$\begin{aligned} h(X) &= -\int_{-\infty}^\infty f_X(x)\left(\ln\frac{1}{2\lambda} - \frac{|x|}{\lambda}\right) dx \\ &= \ln(2\lambda) + \frac{1}{\lambda}\int_{-\infty}^\infty |x| f_X(x) dx \\ &= \ln(2\lambda) + \frac{\lambda}{\lambda} \\ &= 1 + \ln(2\lambda) \text{ nats} \end{aligned}$$

where we have used the fact that $E[|X|] = \lambda$.

(c) X_1 is independent of X_2 and $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$, $i = 1, 2$. Hence, $X = aX_1 + bX_2$ is also Gaussian: $X \sim \mathcal{N}(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2)$. So,

$$h(aX_1 + bX_2) = \frac{1}{2}\ln(2\pi e(a^2\sigma_1^2 + b^2\sigma_2^2)) \text{ nats.}$$

3. (a) Let X be a random variable with pdf $f_X(x)$ and support $\mathcal{X}$. Prove that for any positive C_1,

$$C_1 \cdot E[q(X)] - h(X) \geq \ln(C_2)$$

where

$$C_2 = \frac{1}{\int_{\mathcal{X}} e^{-C_1 \cdot q(X)} dx}.$$

Hint: $D(X\|Y) = \int_{\mathcal{X}} f_X(x) \ln \frac{f_X(x)}{f_Y(x)} dx \geq 0$ for any two continuous random variables X and Y.

(b) When does equality hold in (a) such that

$$h(X) = E[q(X)] - \ln(C) = C_1 \cdot E[q(X)] + \ln\left(\int_{\mathcal{X}} e^{-C_1 \cdot q(x)} dx\right)?$$

(c) Use (a) and (b) to find the random variable that maximizes the differential entropy among all variables with $E[(X-1)^2] = \sigma^2$ and support $\mathbb{R}$.

(d) Use (a) and (b) to find the random variable that maximizes the differential entropy among all variables with finite support $[a, b]$.

Solution.

(a)

$$\begin{aligned}
C_1 \cdot E[q(X)] - h(X) &= C_1 \cdot \int_{-\infty}^{\infty} f_X(x)q(x)dx - \int_{-\infty}^{\infty} f_X(x) \ln \frac{1}{f_X(x)}dx \\
&= \int_{-\infty}^{\infty} f_X(x) \ln \frac{f_X(x)}{e^{-C_1 \cdot q(x)}}dx \\
&= \int_{-\infty}^{\infty} f_X(x) \ln \frac{f_X(x)}{C_2 e^{-C_1 \cdot q(x)}}dx + \ln(C_2) \\
&= \int_{-\infty}^{\infty} f_X(x) \ln \frac{f_X(x)}{f_Y(x)}dx + \ln(C_2) \\
&\geq \ln(C_2)
\end{aligned}$$

where $f_Y(x) = C_2 e^{-C_1 \cdot q(x)}$ is a pdf defined over $x \in \mathcal{X}$.

(b) From the derivation in (a), equality holds iff $f_X(x) = f_Y(y)$, i.e., $f_X(x) = C_2 e^{-C_1 \cdot q(x)}$ for $x \in \mathcal{X}$.

(c) $f_X(x) = \frac{e^{-C_1(x-1)^2}}{\int_{-\infty}^{\infty} e^{-C_1(x-1)^2}dx}$, where C_1 must satisfy

$$E[(X-1)^2] = \int_{-\infty}^{\infty} (x-1)^2 f_X(x)dx = \sigma^2.$$

This implies $C_1 = \frac{1}{2\sigma^2}$ and

$$\begin{aligned}
h(X) &= C_1 \cdot E[(X-1)^2] + \ln\left(\int_{-\infty}^{\infty} e^{-C_1(x-1)^2}dx\right) \\
&= \frac{1}{2} + \frac{1}{2}\ln\left(2\pi\sigma^2\right) = \frac{1}{2}\ln\left(2\pi e\sigma^2\right).
\end{aligned}$$

(d) We can set $q(x) = 1$ (i.e., $E[q(X)] = E[1] = 1$) in (a). Hence, $f_X(x) = \frac{e^{-C_1}}{\int_a^b e^{-C_1}dx} = \frac{1}{b-a}$, where C_1 straightforwardly satisfies

$$E[1] = \int_{-\infty}^{\infty} 1 \cdot f_X(x)dx = 1.$$

As a result, C_1 can be arbitrary, and

$$\begin{aligned} h(X) &= C_1 \cdot E[1] + \ln\left(\int_a^b e^{-C_1} dx\right) \\ &= \ln(b-a). \end{aligned}$$

4. Let the channel model be $Y = AX + Z$, where X and Y are respectively channel input and output, and A and Z are respectively channel gain and additive noise. Suppose A, X and Z are independent and they are all continuous random variables that admit pdf.

   (a) Suppose the channel output can perfectly estimate the value of A. The channel capacity cost function is thus given by

   $$C(P) = \sup_{F_X : E[t(X)] \le P} I(X; Y, A).$$

   In other words, $C(P)$ is the largest input-output mutual information attainable for any input distribution F_X that satisfies the cost constraint $E[t(X)] \le P$. Show that $C(P)$ can be alternatively represented as

   $$C(P) = \left(\sup_{F_X : E[t(X)] \le P} h(Y|A)\right) - h(Z).$$

   Hint: Translation does not change the differential entropy.

   (b) Let $t(x) = x^2$, i.e., $t(\cdot)$ is a power constraint, and let Z be zero-mean Gaussian distributed with variance σ^2. Also assume $E[A^2] = 1$. Is it possible to have

   $$C(P) \stackrel{?}{=} \frac{1}{2} \ln\left(1 + \frac{P}{\sigma^2}\right).$$

   Justify your answer.
   Hint: $E[Y^2] = E[(AX+Z)^2] = E[A^2]E[X^2] + E[Z^2]$

Solution.

(a) By chain rule for mutual information,

$$\begin{aligned} I(X; Y, A) &= I(X; A) + I(X; Y|A) \\ &= I(X; Y|A) \quad (\text{Because } I(X; A) = 0) \\ &= h(Y|A) - h(Y|X, A). \end{aligned}$$

Since $h(Y|X,A) = h(AX+Z|X,A) = h(Z|X,A)$ as differential entropy is translation-invariant, and since $h(Z|X,A) = h(Z)$ as Z is independent of X and A, we have the desired result.

(b) $E[X^2] \le P$ holds iff $E[Y^2] \le P + \sigma^2$, since

$$E[Y^2] = E[(AX+Z)^2] = E[A^2]E[X^2] + E[Z^2] = E[X^2] + \sigma^2.$$

Hence,

$$\begin{aligned} \sup_{F_X : E[X^2] \le P} h(Y|A) &= \sup_{F_X : E[Y^2] \le P + \sigma^2} h(Y|A) \\ &\le \frac{1}{2} \ln \left(2\pi e (P + \sigma^2)\right), \end{aligned}$$

where the inequality follows from that the differential entropy is maximized when Y conditional on A is Gaussian distributed with variance $P+\sigma^2$. However, when $A = a$, we need X to be zero-mean Gaussian distributed with $E[X^2] = P/a$ to achieve the bound, but this requires the channel input (i.e., the transmitter) to have the perfect knowledge about the value of A and we also need $a \ge 1$ to fulfill the constraint $E[X^2] \le P$. As the channel input is not aware of $A = a$,

$$C(P) < \frac{1}{2} \ln \left(1 + \frac{P}{\sigma^2}\right)$$

unless $\Pr[A = 1] = 1$.

Note: So we can only fix X to be Gaussian distributed with $E[X^2] = P$. As such, $E[Y^2|A = a] = a^2 P + \sigma^2$, and $h(Y|A = a) = \frac{1}{2} \ln (2\pi e(a^2 P + \sigma^2))$, which implies

$$h(Y|A) = E_A \left[\frac{1}{2} \ln \left(2\pi e (A^2 P + \sigma^2)\right) \right].$$

Note: When the value A is not known to the receiver but only its distribution, we shall determine the channel capacity via

$$I(X;Y) = h(Y) - h(Y|X) = h(Y) - h(AX+Z|X).$$

In such case, we cannot claim that $h(AX+Z|X) = h(Z)$ and the determination of the exact capacity capacity formula is still an open problem.