Quantum Information Theory
Solutions 8

Exercise 8.1 Fano's Inequality

Given random variables X and Y, how well can we predict X given Y? Fano's inequality bounds the probability of error in terms of the conditional entropy $H(X|Y)$. The goal of this exercise is to prove the inequality

$$P_{\text{error}} \geq \frac{H(X|Y)-1}{\log|X|}.$$

1. *Representing the guess of X by the random variable $\widehat{X}$, which is some function, possibly random, of Y, show that $H(X|\widehat{X}) \geq H(X|Y)$.*

   The random variables X, Y, and $\widehat{X}$ form a Markov chain, so we can use the data processing inequality. It leads directly to $H(X|\widehat{X}) \geq H(X|Y)$.

2. *Consider the indicator random variable E which is 1 if $\widehat{X} \neq X$ and zero otherwise. Using the chain rule we can express the conditional entropy $H(E, X|\widehat{X})$ in two ways:*

   $$H(E, X|\widehat{X}) = H(E|X, \widehat{X}) + H(X|\widehat{X}) = H(X|E, \widehat{X}) + H(E|\widehat{X}) \tag{1}$$

   Calculate each of these four expressions and complete the proof of the Fano inequality. Hint: For $H(E|\widehat{X})$ use the fact that conditioning reduces entropy: $H(E|\widehat{X}) \leq H(E)$. For $H(X|E, \widehat{X})$ consider the cases $E = 0, 1$ individually.

   $H(E|X, \widehat{X}) = 0$ since E is determined from X and $\widehat{X}$. $H(E|\widehat{X}) \leq H(E) = h_2(P_{\text{error}})$ since conditioning reduces entropy.

   $$\begin{aligned} H(X|E, \widehat{X}) &= H(X|E = 0, \widehat{X})p(E = 0) + H(X|E = 1, \widehat{X})p(E = 1) \\ &= 0(1 - P_{\text{error}}) + H(X|E = 1, \widehat{X})P_{\text{error}} \leq P_{\text{error}} \log|X| \end{aligned}$$

   Putting this together we have

   $$H(X|Y) \leq H(X|\widehat{X}) \leq h_2(P_{\text{error}}) + P_{\text{error}} \log|X| \leq 1 + P_{\text{error}} \log|X|,$$

   where the last inequality follows since $h_2(x) \leq 1$. Rearranging terms gives the Fano inequality.

Exercise 8.2 Quantum mutual information

Consider a composed system $A \otimes B \otimes C$ with a shared state ρ_{ABC}.
In a first step we ignore system C and consider only $A \otimes B$ (and the reduced state $\rho_{AB} = Tr_C(\rho_{ABC})$). One way of quantifying the correlations between A and B is to use the mutual information *between them, defined as*

$$\begin{aligned} I(A : B) &= H(A) + H(B) - H(AB) \qquad (2) \\ &= H(A) - H(A|B). \qquad (3) \end{aligned}$$

If we have access to C, we can define a conditional version of the mutual information between A and B as

$$\begin{aligned} I(A : B|C) &= H(A|C) + H(B|C) - H(AB|C) \qquad (4) \\ &= H(A|C) - H(A|BC). \qquad (5) \end{aligned}$$

(a) *Assume a system formed by two qubits A and B that share a state ρ_{AB}. Consider bases $\{|0\rangle_A, |1\rangle_A\}$ and $\{|0\rangle_B, |1\rangle_B\}$ for the subsystems of each qubit.*

1. *Check that the mutual information of the fully entangled state , $|\Psi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, is maximal.*

   The global state is pure and the reduced states on A and B are both fully mixed, $\rho_A = \rho_B = \mathbb{1}/2$, so we have

   $$H(AB) = 0, \quad H(A) = H(B) = 1 \quad \Rightarrow \quad I(A:B) = 2,$$

   which is maximal, because the entropy of a single qubit is at most $\log|\mathcal{H}_A| = 1$, as we saw in exercise 11.2, and the entropy of the joint state is always non negative.

2. *See that for classically correlated states, $\rho_{AB} = p|0\rangle\langle 0|_A \otimes \sigma_B^0 + (1-p)|1\rangle\langle 1|_A \otimes \sigma_B^1$ (where $0 \leq p \leq 1$), the mutual information cannot be greater than one.*

   We can rewrite the mutual information as

   $$I(A:B) = \underbrace{H(A)}_{\leq 1} - \underbrace{H(A|B)}_{\geq 0^{(*)}} \leq 1$$

   where $^{(*)}$ comes from exercise $11.1.b)3$.

(b) *Consider the so-called* cat state *shared by four qubits, $A \otimes B \otimes C \otimes D$, that is defined as*

$$|☺\rangle = \frac{1}{\sqrt{2}}(|0000\rangle + |1111\rangle). \tag{6}$$

Check how the mutual information between qubits A and B changes with the knowledge of the remaining qubits, namely:

1. $I(A:B) = 1$.
2. $I(A:B|C) = 0$.
3. $I(A:B|CD) = 1$.

The reduced states of the system for k qubits (which are independent of the qubits traced out) have entropies denoted by h_k, given as follows:

$$\begin{aligned}
\rho_4 &= |☺\rangle\langle☺| && \Rightarrow h_4 = 0, \\
\rho_3 &= \frac{1}{2}(|000\rangle\langle 000| + |111\rangle\langle 111|) && \Rightarrow h_3 = 1, \\
\rho_2 &= \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|) && \Rightarrow h_2 = 1, \\
\rho_1 &= \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) && \Rightarrow h_1 = 1.
\end{aligned}$$

The mutual information between A and B given the knowledge of other qubits comes

$$\begin{aligned}
I(A:B) &= H(A) + H(B) - H(AB) \\
&= h_1 + h_1 - h_2 = 1, \\
I(A:B|C) &= H(A|C) + H(B|C) - H(AB|C) \\
&= H(AC) - H(C) + H(BC) - H(C) - H(ABC) + H(C) \\
&= h_2 - h_1 + h_2 - h_1 - h_3 + h_1 = 0, \\
I(A:B|CD) &= H(A|CD) + H(B|CD) - H(AB|CD) \\
&= H(ACD) - H(CD) + H(BCD) - H(CD) - H(ABCD) + H(CD) \\
&= h_3 - h_2 + h_3 - h_2 - h_4 + h_2 = 1.
\end{aligned}$$

Exercise 8.3 Mutual Information - Weather bet

a) *Compute the mutual information between your guess and the actual weather, and do the same for your grandfather. Remember that all your grandfather knows is that it rains on 80% of the days. You know that as well and you also listen to the weather forecast and know that it is right 80% of the time and is always correct when it predicts rain.*

The mutual information is given by $I(X:Y) = H(X) - H(X|Y)$. Let us call your grandfather G, you Y and the actual weather W. We may also assume you followed the radio forecast (which we saw was an optimal strategy) and we will hold on to the notation $\hat{R}, \hat{S}$ for guesses (both yours and your grandfather's). Then we have, for the grandfather

$$\begin{aligned}
H(W) &= -P(R)\log P(R) - P(S)\log P(S) \\
&= -0.8\log 0.8 - 0.2\log 0.2 \\
H(G) &= -P(\hat{R})\log P(\hat{R}) - P(\hat{S})\log P(\hat{S}) \\
&= -1.\log 1. - 0. = 0 \\
H(GW) &= -P(\hat{R}R)\log P(\hat{R}R) - P(\hat{S}R)\log P(\hat{S}R) - P(\hat{R}S)\log P(\hat{R}S) - P(\hat{S}S)\log P(\hat{S}S) \\
&= -0.8\log 0.8 - 0 - 0 - 0.2\log 0.2 \\
H(W|G) &= H(GW) - H(G) \\
&= -0.8\log 0.8 - 0.2\log 0.2 \\
I(W:G) &= H(W) - H(H|G) \\
&= 0.
\end{aligned}$$

For your case we will calculate the conditional entropy directly,

$$\begin{aligned}
H(W) &= -0.8\log 0.8 - 0.2\log 0.2 \\
H(W|Y) &= -P(\hat{R})\left[P(R|\hat{R})\log P(R|\hat{R}) + P(S|\hat{R})\log P(S|\hat{R})\right] \\
&\quad - P(\hat{S})\left[P(R|\hat{S})\log P(R|\hat{S}) + P(S|\hat{S})\log P(S|\hat{S})\right] \\
&= -0.6[1\log 1 + 0] - 0.4[0.5\log 0.5 + 0.5\log 0.5] \\
&= -0.4\log 0.5 \\
I(W:Y) &= H(W) - H(H|Y) \\
&= -0.8\log 0.8 - 0.2\log 0.2 + 0.4\log 0.5 \\
&= 0.32.
\end{aligned}$$

b) *What would your strategy be? And your grandfather's? After N days, what is the expected gain for each of you? What is the probability that he finishes with more money than you?*

Here your optimal strategy and your grandfather's optimal strategy are the strategies that maximize the total amount of money you will each have after N days.

Let us first compute your grandfather's expected gain after N days. If your grandfather bets all his money on rain every evening (since this event is most likely for him each day), his expected amount of money is:

$$\langle \pounds_G \rangle = 0.8^N 2^N = 1.6^N.$$

As for you, how much you bet when the radio predicts a sunny day has no impact on your expected gain (because you will guess correctly 50% of the time: $P_{R|\hat{S}} = P_{S|\hat{S}} = 50\%$). So you should choose to keep your money when the forecast is sun (so that you don't lose with probability 50%). Your expected amount of money is

$$\begin{aligned} \langle \pounds_Y \rangle &= \sum_{k=0}^{N} P_{\hat{R}}(k)\ 2^k \\ &= \sum_{k=0}^{N} \binom{N}{k} 0.6^k\ 0.4^{N-k}\ 2^k \\ &= 1.6^N, \end{aligned}$$

where $P_{\hat{R}}(k)$ is the probability that the radio predicts rain on exactly k of the N days.

The expected amount of money for both of you is the same!

If there is rain every day then $k = N$, and your grandfather will have 2^N pounds. What is the probability that you have less money than your grandfather in this case? You will have to have a prediction of rain for each day. This is $P_{\hat{R}|R}^N = 0.75^N$. Therefore the probability that your grandfather makes more money than you when $k = N$ is $1 - (0.75)^N$.

If there is one or more days of sun ($k \neq N$), then your grandfather will have lost all his money. You will never lose all your money (either you don't bet when the forecast is sunny, or you are 100% sure that when you bet all your money on rain when the forecast is rain).

Therefore, your grandfather has more money than you with probability $1-0.8^N(1-0.75^N)$.

In other words, if you play this game for many days, you're extremely likely to have more money than your grandfather.

If you chose to bet randomly when the radio predicts sun, then the above probability of having more than your grandfather will decrease, because you could lose all your money with this strategy.