ECE 788: Network Information Theory
Assignment 2 (due on Sept. 21)

1. Consider alphabets $\mathcal{X} = \mathcal{Y} = \{0, 1\}$, $n = 12$ and the following joint types

$$P_{XY}^{(1)}\colon \begin{bmatrix} 1/6 & 1/3 \\ 1/3 & 1/6 \end{bmatrix},\ P_{XY}^{(2)}\colon \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix} \text{ and } P_{XY}^{(3)}\colon \begin{bmatrix} 1/4 & 1/4 \\ 1/4 & 1/4 \end{bmatrix}.$$

Fix a sequence $x^{12} = (010111110000)$.
1.a. Check that such sequence has a marginal type $N(a|x^n)/n$ that is consistent with the three joint types above.
1.b. Calculate the conditional entropy $H(Y|X)$ (a.k.a. $H(P_{Y|X})$) for the three joint distributions above. For $P_{XY}^{(2)}$, how would you interpret this result (in terms of uncertainty)? For $P_{XY}^{(3)}$, compare $H(P_{Y|X})$ with $H(P_Y)$: Why are they the same (provide an intuitive explanation in terms of uncertainty)?
1.c. From the point above, guess the size of the set

$$T_0^{12}(P_{XY}^{(k)}|x^{12}) = \{y^{12} \in \mathcal{Y}^{12}\colon (x^{12}, y^{12}) \in P_{XY}^{(k)}\}$$

of sequences y^{12} that have the given joint type $P_{XY}^{(k)}$ with x^{12} for $k = 1, 2, 3$. For what joint type you expect the size of this set to be larger?

2. Using the weak law of large numbers, we want to prove the following simplified statement of the result (1.12) (see also (5) in the handout) of the AEP:

$$\lim_{n\to\infty} \Pr\left[X^n \in T_\epsilon^n(P_X)\right] = 1,$$

where X^n is i.i.d. with probability P_X.
2.a. To do this, first argue that the above is the same as proving that

$$\lim_{n\to\infty} \Pr\left[\left|\frac{N(a|x^n)}{n} - P_X(a)\right| > \epsilon P_X(a)\right] = 0$$

for all $a \in \mathcal{X}$. If $P_X(a) = 0$, the above is obvious (why?), so we need to consider $P_X(a) > 0$.
2.b. With $P_X(a) > 0$, write $N(a|x^n)$ as a sum of independent random variables (Hint: The indicator function $1(x_i = a)$ that equals 1 if $x_i = a$ and zero otherwise may come in handy)
2.c. Having completed the above, apply the weak law of large numbers and complete the proof (Even better, complete the proof directly using the Markov inequality).
Remark: A proof of (1.12) ((5) in the handout) requires the use of Chernoff bounds, which can be obtained from the Markov inequality as shown on p. 168-169 (see booklet).

3. Let $x^n \in T_\epsilon^n(P_X)$ and $g(x)$ a non-negative function.
3.a. Prove that

$$(1-\epsilon)E[g(X)] \le \frac{1}{n}\sum_{i=1}^{n} g(x_i) \le (1+\epsilon)E[g(X)],$$

where the expectation is taken with respect to P_X (Hint: Rewrite $\frac{1}{n}\sum_{i=1}^{n} g(x_i)$ as a sum over $a \in \mathcal{X}$ and then use the definition of type and of expectation).
3.b. Set $g(x) = 1(x = a)$ for some $a \in \mathcal{X}$. What do we get? (Hint: $E[1(X_i = a)] = P_X(a)$).