ECE642 Information Theory and Source Coding — Mark R. Bell
Fall 2023 — MSEE 336

Homework Assignment #1

Reading Assignment: Ch. 1; Ch. 2: 2.1–2.8, 2.10

Due Friday September 8, 2023

1. Two fair dice are thrown. The outcome of the experiment is taken as the sum of the resulting outcomes of the two dice. What is the entropy of this experiment?
2. Suppose there are 3 urns, each containing two balls. The first urn contains two black balls; the second contains one black ball and one white ball; the third contains two white balls. An urn is selected at random (prob. 1/3 of selecting each urn), and one ball is removed from this urn. Let $X = \{1, 2, 3\}$ represent the sample space of the urn selected, and let $Y = \{B, W\}$ be the sample space of the color ball selected. Calculate the entropies $H(X)$, $H(Y)$, $H(XY)$, $H(X|Y)$, $H(Y|X)$, and the mutual information $I(X;Y)$.
3. Consider a horse race in which there are a total of 7 horses, 4 black horses and 3 grey horses; hence the sample space of the horse race can be represented by $X = \{b_1, b_2, b_3, b_4, g_5, g_6, g_7\}$. Subdivide the horse race such that we have two experiments, $Y = \{b_1, \ldots, b_4\}$, the outcome among the black horses, and $Z = \{g_5, \ldots, g_7\}$ the outcome among the grey horses. Define the event B as the event that a black horse wins. Assuming that in the original experiment, the *a priori* probabilities of any particular horse winning is equal, verify that
$$H(X) = H(B) + P(B)H(Y) + P(\overline{B})H(Z).$$
4. Let X be a random variable that only assumes *non-negative integer* values. Suppose $\mathrm{E}[X] = M$, where M is some fixed number. In terms of M, how large can $H(X)$ be? Describe the extremal X as completely as you can.
5. In class, we stated the following theorem for convex functions:

   Theorem. *If $\lambda_1, \ldots, \lambda_N$ are non-negative numbers whose sum is unity, then for every set of points $\{P_1, \ldots, P_N\}$ in the domain of the convex $\cap$ function $f(P)$, the following inequality is valid:*
   $$f\left(\sum_{n=1}^{N} \lambda_n P_n\right) \geq \sum_{n=1}^{N} \lambda_n f(P_n).$$
   Prove this theorem, and using this result, prove Jensen's inequality:

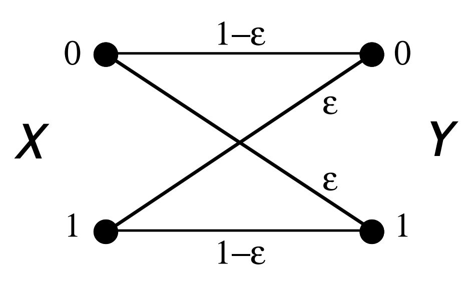
   Jensen's Inequality. *Assume X is a random variable taking on values from the set $R_X = \{x_1, \ldots, x_N\}$ with probabilities $p_1, \ldots, p_N$. If $f(x)$ be a convex $\cap$ function whose domain includes R_X, then*
   $$f(\mathrm{E}[X]) \geq \mathrm{E}[f(X)].$$
   Furthermore, if $f(x)$ is strictly convex, equality holds if and only if one particular element of R_X is certain.
   (Hint: The initial theorem is most easily proven by induction.)

6. In class, we showed that if X and Y are two *independent* jointly-distributed discrete random variables with probability mass function (pmf) $p(x,y)$, then the joint entropy $H(X,Y)$ is given by
$$H(X,Y) = H(X) + H(Y).$$
We now wish to generalize this result. Show that if X and Y are two jointly-distributed discrete random variables with pmf $p(x,y)$,, that in general
$$H(X,Y) \leq H(X) + H(Y),$$
with equality if and only if the outcomes from the marginal ensembles X and Y are statistically independent.

7. Prove the following three equalities of (average) mutual information that were stated in class:
   (i) $I(X;Y) = I(Y;X)$,
   (ii) $I(X;Y) = H(Y) - H(Y|X)$,
   (iii) $I(X;Y) = H(X) + H(Y) - H(X,Y)$.

8. Let $\mathbf{X}$ be a discrete random variable taking on values $2, 3, 4, \ldots$, with probability
$$P\{\mathbf{X} = n\} = \frac{1}{An \log^2 n},$$
where A is a constant. Show that $H(\mathbf{X}) = +\infty$, that is, that the defining series for $H(\mathbf{X})$ does not converge.

9. Consider a binary symmetric channel with *error crossover probability* ϵ as shown below. Assume the input ensemble is $X = \{0, 1\}$ with $P_X(0) = P_X(1) = 1/2$. Calculate $I(X;Y)$ in this situation as a function of ϵ.



10. *Cover and Thomas:* Problem 2.17. (Problem 2.7 in first edition.)