Information and Coding Theory: Exercise Sheet 1

Task 1. Prove Jensen's inequality: if f is convex, and X is a random variable with a finite set of values, then

$$\mathbb{E}[f(x)] \geq f(\mathbb{E}[X]) \ .$$

Show that if f is strictly convex, then the equality holds if and only if X is a constant function.

Task 2. Prove Log-sum inequality: for $a_1, \ldots, a_n, b_1, \ldots, b_n > 0$

$$\sum a_i \log \frac{a_i}{b_i} \geq \sum a_i \log \frac{\sum_j a_j}{\sum_j b_j}$$

Show that the equality holds if and only if $\frac{a_i}{b_i}$ has the same value for all i.

Task 3. Prove Gibbs inequality:

Let $p_1, \ldots, p_n$ be a probability ditribution and $q_1, \ldots, q_n > 0$ satisfy $\sum_i q_i \leq 1$. Then

$$-\sum_i p_i \log p_i \leq -\sum_i p_i \log q_i .$$

Task 4. Prove general properties of entropy (you may use all inequalities given above, even if you cannot prove them):

$$\begin{aligned}
&& 0 \leq H[X] \leq \log|U| && \text{where } X : \Omega \to U \\
&& 0 \leq H[X|Y] \leq H[X] && \\
X, Y \text{ are independent} &\implies& H[Y|X] = H[Y] && \\
X, Y \text{ are independent} &\implies& H[Y, X] = H[X] + H[Y] && \\
&& H[X|Y] \leq H[X, Y] \leq H[X] + H[Y] && \\
&& 0 \leq I(X; Y) \leq \min(H[X], H[Y]) && \\
&& I(X; Y) = H[X] + H[Y] - H[X, Y] &&
\end{aligned}$$

Task 5. Show that for any random variables X, Y we have:

$$H[Y|X] = 0 \iff Y = f(X)$$

Use this to conclude that for any function f we have

$$H[f(X)] \leq H[X] \ .$$

Task 6. Prove general chain rule: for random variables $X_1, \ldots, X_n$:

$$H[X_1, \ldots, X_n] = H[X_1] + \sum_{i=2}^{n} H[X_i | X_1, \ldots, X_{i-1}] \ .$$

Task 7. Suppose that you have n balls, one of them is false and has different weight. In one comparison you can compare two sets of balls and the answer is "equal"/"non-equal".

Use entropy and conditional entropy to lower bound the needed number of comparisons. Consider the variable B that gives the number of false ball and random variables that represent the consecutive comparisons. Note, the second comparison may depend on the first, so you should not directly use its entropy. However, when the first comparison is known, the second has only two possible outcomes. You can use Task 6 even if you cannot show it.

Use a similar approach to the case when there are three possible answers (lighter, equal, heavier) and we also want to know, whether the false ball is lighter or heavier.

Task 8 (Grouping property)**.** Let $p = p_1, \ldots, p_n$ be a probability distribution. Define $h(p_1, \ldots, p_n) = -\sum_{i=1}^{n} p_i \log p_i$, i.e. as entropy for random variable, but for probability distribution.

Show that

$$h(p_1, \ldots, p_n) = h(p_1 + p_2, p_3, \ldots, p_n) + (p_1 + p_2) h\left(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2}\right)$$

Task 9 (Entropy of a sum). Let X, Y be random variables (with finite set of values) and $Z = X + Y$. Show that

1. $H[Z|X] = H[Y|X]$ and that when X, Y are independent, then $H(Y) \leq H[Z]$ and $H[X] \leq H[Z]$.
2. Under what conditions does $H[Z] = H[X] + H[Y]$ hold?