Homework 1 Solutions

13) **Inequality.** Show $\ln x \geq 1 - \frac{1}{x}$ for $x > 0$.

13) *Inequality*. Using the Remainder form of the Taylor expansion of $\ln(x)$ about $x = 1$, we have for some c between 1 and x

$$\ln(x) = \ln(1) + \left(\frac{1}{t}\right)_{t=1}(x-1) + \left(\frac{-1}{t^2}\right)_{t=c}\frac{(x-1)^2}{2} \leq x - 1$$

since the second term is always negative. Hence letting $y = 1/x$, we obtain

$$-\ln y \leq \frac{1}{y} - 1$$

or

$$\ln y \geq 1 - \frac{1}{y}$$

with equality iff $y = 1$.

19) **Infinite entropy.** This problem shows that the entropy of a discrete random variable can be infinite. Let $A = \sum_{n=2}^{\infty}(n\log^2 n)^{-1}$. (It is easy to show that A is finite by bounding the infinite sum by the integral of $(x\log^2 x)^{-1}$.) Show that the integer-valued random variable X defined by $\Pr(X = n) = (An\log^2 n)^{-1}$ for $n = 2, 3, \ldots$, has $H(X) = +\infty$.

Infinite entropy. By definition, $p_n = \Pr(X = n) = 1/An\log^2 n$ for $n \geq 2$. Therefore

$$\begin{aligned} H(X) &= -\sum_{n=2}^{\infty} p(n)\log p(n) \\ &= -\sum_{n=2}^{\infty}\left(1/An\log^2 n\right)\log\left(1/An\log^2 n\right) \\ &= \sum_{n=2}^{\infty}\frac{\log(An\log^2 n)}{An\log^2 n} \\ &= \sum_{n=2}^{\infty}\frac{\log A + \log n + 2\log\log n}{An\log^2 n} \\ &= \log A + \sum_{n=2}^{\infty}\frac{1}{An\log n} + \sum_{n=2}^{\infty}\frac{2\log\log n}{An\log^2 n}. \end{aligned}$$

The first term is finite. For base 2 logarithms, all the elements in the sum in the last term are nonnegative. (For any other base, the terms of the last sum eventually all become positive.) So all we have to do is bound the middle sum, which we do by comparing with an integral.

$$\sum_{n=2}^{\infty}\frac{1}{An\log n} > \int_2^{\infty}\frac{1}{Ax\log x}dx = K\ln\ln x\Big|_2^{\infty} = +\infty.$$

We conclude that $H(X) = +\infty$.

21) **Markov's inequality for probabilities.** Let $p(x)$ be a probability mass function. Prove, for all $d \geq 0$,

$$\Pr\{p(X) \leq d\}\ \log\left(\frac{1}{d}\right) \leq H(X). \tag{8}$$

Markov inequality applied to entropy.

$$\begin{aligned}
P(p(X) < d) \log \frac{1}{d} &= \sum_{x:p(x)<d} p(x) \log \frac{1}{d} && (47) \\
&\leq \sum_{x:p(x)<d} p(x) \log \frac{1}{p(x)} && (48) \\
&\leq \sum_{x} p(x) \log \frac{1}{p(x)} && (49) \\
&= H(X) && (50)
\end{aligned}$$

Problem 4. Read the statement and proof of Han's inequality from Polyansky and Wu, page 16. Write the statement of the inequality and prove it on your own after the initial reading. Please do not copy the text directly, try to explain things your way.

Problem 5. Three squares have average area $\bar{a} = 100\text{m}^2$. The average of the lengths of their sides is $\bar{l} = 10\text{m}$. What can be said about the area of the larges square?

Solution: Let x be the length of the side of a square, and let the probability of x be $(1/3, 1/3, 1/3)$ over the three lengths (l_1, l_2, l_3). Then the information that we have is that $E[x] = 10$ and $E[f(x)] = 100$, where $f(x) = x^2$ is the function mapping lengths to areas. This is a strictly convex function. We notice that the equality $E[f(x)] = f(E[x])$ holds, therefore x is a constant, and the three lengths must all be equal. The area of the largest square is $100m^2$.

Problem 6. The Rényi entropy of order $\alpha \geqslant 0$, $\alpha \neq 1$ of a discrete RV X supported on a set of cardinality M is defined as

$$H_\alpha(X) = \frac{1}{1-\alpha} \log \left(\sum_i^M p_i^\alpha \right).$$

1. Show that $H_0 \geqslant H_1 \geqslant H_2 \geqslant H_\infty$. Observe that the subscripts 1 and ∞ are to be taken in the sense of a limit (i.e., $\alpha \to 1$, $\alpha \to \infty$, respectively).
2. Show that Rényi entropy is nonnegative, and that it is concave for $\alpha \leqslant 1$.

1 (10 points total) 2.5 pts each: inequality 1, inequality 2, inequality 3, taken in sense of limit or find derivative w.r.t. alpha and explain

2 (10 points total) non-negative (5 pts), concavity (5pts)

Proof. For $0 \leq \alpha < \beta$ with $\alpha \neq 1$ and $\beta \neq 1$,

$$\begin{aligned}
H_\alpha(X) &= \frac{1}{1-\alpha} \log \sum_{x \in \mathcal{X}} P_X(x)^\alpha \\
&= -\log \mathrm{E}\Big[P_X(X)^{\alpha-1}\Big]^{\frac{1}{\alpha-1}} \\
&= -\log \mathrm{E}\Big[P_X(X)^{\alpha-1}\Big]^{\frac{\beta-1}{\alpha-1}\frac{1}{\beta-1}} \\
&\geq -\log \mathrm{E}\Big[P_X(X)^{(\alpha-1)\frac{\beta-1}{\alpha-1}}\Big]^{\frac{1}{\beta-1}} \\
&= -\log \mathrm{E}\Big[P_X(X)^{\beta-1}\Big]^{\frac{1}{\beta-1}} \\
&= \frac{1}{1-\beta} \log \sum_{x \in \mathcal{X}} P_X(x)^\beta \\
&= H_\beta(X).
\end{aligned}$$

We observe that x^c is convex-$\cup$ for $c \geq 1$ or $c \leq 0$ and convex-$\cap$ for $0 \leq c \leq 1$. The inequality in the above derivation follows from the Jensen inequality in the following cases:

$\beta > \alpha > 1 :\ c = \frac{\beta-1}{\alpha-1} > 1$, x^c is convex-$\cup$ and $\frac{1}{\beta-1} > 0$;

$\beta > 1 > \alpha \geq 0 :\ c = \frac{\beta-1}{\alpha-1} < 0$, x^c is convex-$\cup$ and $\frac{1}{\beta-1} > 0$;

$1 > \beta > \alpha \geq 0 :\ 1 > c = \frac{\beta-1}{\alpha-1} > 0$, x^c is convex-$\cap$ and $\frac{1}{\beta-1} < 0$.

For $\alpha = 1$ or $\beta = 1$, the Jensen inequality can be applied directly. The conditions for equality in (2.14) follow directly from the Jensen inequality. $\square$

$$H_\alpha(\lambda P+(1-\lambda)Q)=\frac{1}{1-\alpha}\log\left[\Sigma(\lambda p_i+(1-\lambda)q_i)^\alpha\right] \quad (14)$$

$$>\frac{1}{1-\alpha}\log\left[\Sigma(\lambda p_i^\alpha+(1-\lambda)q_i^\alpha)\right] \quad (15)$$

$$=\frac{1}{1-\alpha}\log\left[\lambda\Sigma p_i^\alpha+(1-\lambda)\Sigma q_i^\alpha\right]. \quad (16)$$

Since the log function is concave, (16) is strictly greater than

$$\frac{1}{1-\alpha}\left[\lambda \log(\Sigma p_i^\alpha)+(1-\lambda)\log(\Sigma q_i^\alpha)\right]$$

$$=\lambda H_\alpha(P)+(1-\lambda)H_\alpha(Q) \quad (17)$$

(Solution Sources: TC Solution Manual, Aarti Singh Lecture notes, ECE 534 UIC notes, Renyi's entropy and probability of error, Ben-Bassat, Raviv, Cachin PhD dissertation.)