Learner's Space 2025
Information Theory and Coding
Assignment 1

Suketu Patni

June 15, 2025

1. Find the entropy of a random variable which has a geometric distribution of parameter $p \in (0, 1]$. Express it in terms of $H_2(p)$. [2 marks]

2. a. Find the entropy of a random variable which has a Poisson distribution of parameter $\lambda > 0$. Express it in terms of $f(\lambda)$, where

$$f(\lambda) = \sum_{k=2}^{\infty} \frac{\lambda^k \ln(k!)}{k!}$$

   [3 marks]

   b. Prove that the resultant entropy is finite by proving that $f(\lambda)$ converges for all $\lambda \in \mathbb{R}^+$. (Hint: use the ratio test (read it up if you are unaware of it!)) [2 marks]

3. A discrete random variable can have infinite entropy. Work through the below steps to show an example for the same.

   a. Use the condensation test (you know what to do if you don't know about it!) and the fact that

$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$$

   converges iff $\alpha > 1$ to show that that the series

$$\sum_{n=2}^{\infty} \frac{1}{n(\ln n)^\beta}$$

   converges iff $\beta > 1$. [2 marks]

   b. Put

$$\gamma = \sum_{n=2}^{\infty} \frac{1}{n(\ln n)^2}$$

which exists by the previous part. Show that the random variable X with $\operatorname{supp} X = \mathbb{N}$ and

$$P(X = n) = \frac{1}{\gamma(n+1)(\ln(n+1))^2}$$

has infinite entropy. [4 marks]

4. (just for fun) Solve exercise 2.39 of Mackay's book on information theory programmatically. You need only submit the final numerical answer upto 5 digits after the decimal point. [1 mark]

5. **Typical Sets**: They are a very common construct in information theoretic arguments and form one of the most important applications of the AEP.
   Let X be a random variable with finite image $\mathcal{X}$. The AEP states that for i.i.d random variables $X_1, X_2, \ldots, X_n$ all distributed as X,

   $$-\frac{1}{n}\log_2\left(p(X_1, X_2, \ldots, X_n)\right) \xrightarrow{a.s.} H(X)$$

   So we see that $p(X_1, X_2, \ldots, X_n)$ starts to approach $2^{-nH(X)}$. The typical set $A_{\epsilon,n}$ is the subset of $\mathcal{X}^n$ with probability "very close" to $2^{-nH(X)}$. More precisely,

   $$A_{\epsilon,n} = \left\{(x_1, x_2, \ldots, x_n) \in \mathcal{X}^n : p(x_1, x_2, \ldots, x_n) \in \left[2^{-n(H(X)+\epsilon)}, 2^{-n(H(X)-\epsilon)}\right]\right\}$$

   where $p(\cdot, \cdot, \ldots, \cdot)$ is of course the joint pmf of $X_1, X_2, \ldots, X_n$.
   Prove that

   (a) $P(A_{\epsilon,n}) > 1 - \epsilon \; \forall n > N$ for some $N \in \mathbb{N}$. [2 marks]

   (b) $|A_{\epsilon,n}| \in \left[(1-\epsilon)2^{n(H(X)-\epsilon)}, 2^{n(H(X)+\epsilon)}\right] \; \forall n > M$ for some $M \in \mathbb{N}$. [4 marks]