B3 - Information Theory : Back-paper Exam

Yogeshwaran D.

June 5, 2026. **Time : 2.00 - 5.00 PM.** **Max. points : 50.**

Answer all questions. All questions carry 10 points.
Give necessary justifications and explanations for all your arguments. If you are citing results from the class, mention it clearly.

1. Consider an additive channel (without feedback) whose input alphabet is $\mathcal{X} = \{-2, -1, 0, 1, 2\}$ and whose output is $Y = X + Z$ where Z is distributed uniformly over the interval $[-1,]$. Calculate the capacity of this channel.

2. Let X be an Exponential(1) random variable and Z be an independent standard normal random variable. For all $a \in \mathbb{R}, b > 0$, compute entropies of $bX, a + bZ, a + b\sqrt{2X}Z$.

3. Below we give two examples of channels $(\mathcal{X}, p(y \mid x), \mathcal{X})$ and source message from $\mathcal{W}$. Find the rate of transmission and channel capacity in each of the example. Also in both examples find encoding scheme $\mathcal{W} \to \mathcal{X}$ and decoding scheme $\mathcal{X} \to \mathcal{W}$ such that the probability of error for a uniformly trasmitted word W from $\mathcal{W}$ is at most 0.2 i.e., $\mathbb{P}(\hat{W} \neq W) \leq 0.2$.
   (a) Consider Binary Symmetric Channel on $\mathcal{X} = \{0, 1\}^3$ with crossover (error) probability $\alpha = 0.1$ for each bit i.e., each bit is independently flipped with probability α. Let $\mathcal{W} = [4]$ i.e., W is a uniformly chosen message from $[4]$. **(5)**
   (b) Consider Binary Symmetric Channel on $\mathcal{X} = \{0, 1\}^3$ with erasure probability $\alpha = 0.1$ for each bit i.e., each bit is independently erased with probability α. Let $\mathcal{W} = [8]$ i.e., W is a uniformly chosen message from $[8]$. **(5)**

4. Suppose $\mathcal{X} = \{1, 2, 3, 4\}$, $\hat{\mathcal{X}} = \{1, 2, 3, 4\}$ and $X_1, X_2, \ldots$ are i.i.d. $\sim p(x)$, a pmf on $[4]$. The distortion matrix $d(x, \hat{x})$ is given by

$x \backslash \hat{x}$	1	2	3	4
1	0	0	1	1
2	0	0	1	1
3	1	1	0	0
4	1	1	0	0

   Find $R(D)$ for all $D \geq 0$.

5. Let P and Q be distributions on a finite set $\mathcal{X}$. Let $E \subseteq \mathcal{X}^n$ be such that $P^n(E) = \frac{1}{2}$. What is $\lim_{n\to\infty} \frac{1}{n} \log Q^n(E)$? (You just need to write the expression in terms of P and Q.)

COURSE RECAP

Relative entropy/ KL Divergence: $D(p\|q) = \sum_x p(x) \log \frac{p(x)}{q(x)}$.
Mutual information: $I(X;Y) = D(p_{XY}\|p_X p_Y)$.

Let $X_1, \dots, X_n$ be i.i.d. $\sim p$. **Typical set.** For $\epsilon > 0$, $A_\epsilon^{(n)} = \left\{x^n : \left|-\frac{1}{n}\log p(x^n) - H(X)\right| \leq \epsilon\right\}$.

- $\Pr\{X^n \in A_\epsilon^{(n)}\} \to 1$; $|A_\epsilon^{(n)}| \leq 2^{n(H+\epsilon)}$.
- **Source Coding theorem:** For every $\epsilon > 0$, there exists encoding and decoding $\mathcal{X}^n \to \{0,1\}^{n(H+\epsilon)} \to \mathcal{X}^n$ such the error probability (i.e., $\mathbb{P}(X^n \neq \hat{X}^n)$) vanishes asymptotically.

Optimal Code Length

- A source code maps symbols x to binary strings of length $l(x)$.
- Optimal code length - $L^* := \min\{\sum_x l(x)p(x) \ : \ \sum_x 2^{-l(x)} \leq 1, \, l(x) \in \mathbb{N}\}$.
- $H(X) \leq L^* < H(X) + 1$,
- Shannon coding: $l(x) = \lceil -\log p(x) \rceil$; Huffman coding achieves optimal prefix codes

A **discrete memoryless channel** is specified by $(\mathcal{X}, p(y \mid x), \mathcal{Y})$, where $p(\cdot \mid \cdot)$ is a transition matrix. **Capacity.** $C = \max_{p_X(x)} I(X;Y)$.

Channel Coding Theorem (with and without feedback.) If $R < C$, there exist codes with $\lambda^{(n)} \to 0$. If $R > C$, then $P_e^{(n)} \not\to 0$ for any code

Rate Distortion theory: $R(D) = \min_{p(\hat{x}|x)}\{I(X;\hat{X}) : \mathbb{E}\Big(d(X,\hat{X})\Big) \leq D\}$.
$D(R) = \min_{p(\hat{x}|x)}\{\mathbb{E}\Big(d(X,\hat{X})\Big) : I(X;\hat{X}) \leq R\}$

For $R > R(D)$, there exist $(2^{nR}, n)$-code achieving distortion D and conversely, if there exist $(2^{nR}, n)$-code achieving distortion D then $R \geq R(D)$.

For Binary symmetric channel with input Ber(p), $R(D) = H(\text{Ber}(p)) - H(\text{Ber}(D))$ for $D \leq \min\{p, 1-p\}$ and else 0.

Sanov's theorem: Let X^n be i.i.d $\sim Q$ on $\mathcal{X}$. If $\overline{\mathring{E}} = E$ then $n^{-1}Q^n(E) \to -D(P^* \parallel Q)$.