

MTHE 474 Notes

Timothy Liu

September 15, 2022

Chapter 2

Information Measures for Discrete Systems

Definitions

- **Definition 2.2:** Entropy of discrete random variable X with pmf $P_X(*)$ is defined as

$$H(x) := - \sum_{x \in X} P_X(x) * \log_2 P_X(x)$$

- **Definition 2.2:** Entropy of discrete random variable X with pmf $P_X(*)$ is defined as

$$H(x) := - \sum_{x \in X} P_X(x) * \log_2 P_X(x)$$

- **Definition 2.8 (Joint entropy):**

$$H(X, Y) := - \sum_{(x, y) \in \mathcal{X} \times \mathcal{Y}} P_{X, Y}(x, y) * \log_2 P_{X, Y}(x, y)$$

- **Definition 2.9 (Conditional entropy):**

$$H(Y|X) := \sum_{x \in \mathcal{X}} P_X(x) \left(- \sum_{y \in \mathcal{Y}} P_{Y|X}(y|x) * \log_2 P_{Y|X}(y|x) \right)$$

•

Lemmas/Theorems

- **Lemma 2.4 (Fundamental Inequality):** $\forall x > 0$ and $D > 1$ we have

$$\log_D(x) \leq \log_D e * (x - 1)$$

- **Lemma 2.5 (Non-negativity):** $H(X) \geq 0$

- **Lemma 2.6 (Entropy Upper-Bound):** $H(X) \leq \log_2 |\mathcal{X}|$ where random variable X takes values from finite set \mathcal{X}

- **Lemma 2.7 (Log-Sum inequality)** Write this one out later

- **Theorem 2.10 (Chain rule for entropy):** $H(X, Y) = H(X) + H(Y|X)$

Mutual Information

Definitions

- **Definition 2.2.1 (Mutual Information):**

$$I(X; Y) := H(X) - H(X|Y)$$

- **Definition 2.2.2 (Conditional Mutual Information):**

$$I(X; Y|Z) := H(X|Z) - H(X|Y, Z)$$

Lemmas

- **Lemma 2.15 (Properties of Mutual Information):**

$$1. I(X; Y) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} P_{X, Y}(x, y) \log_2 \frac{P_{X, Y}(x, y)}{P_X(x) P_Y(y)} \quad (1)$$

$$2. I(X; Y) = I(Y; X) = H(Y) - H(Y|X) \quad (2)$$

$$3. I(X; Y) = H(X) + H(Y) - H(X, Y) \text{ equality iff } X \text{ is a function of } Y \quad (3)$$

$$4. I(X; Y) \leq H(X) \quad (4)$$

$$5. I(X; Y) \leq 0 \text{ with equality iff } X \text{ and } Y \text{ are independent} \quad (5)$$

$$6. I(X; Y) \leq \min\{\log_2 |\mathcal{X}|, \log_2 |\mathcal{Y}|\} \quad (6)$$

- **Lemma 2.16 (Chain Rule for Mutual Information):**

$$I(X; Y, Z) = I(X; Y) + I(X; Z|Y) = I(X; Z) + I(X; Y|Z)$$

- **Theorem 2.17 (Chain Rule for entropy):** $X^n := (X_1, \dots, X_n)$ and $x^n := (x_1, \dots, x_n)$

$$H(X^n) = \sum_{i=1}^n H(X_i | X^{i-1})$$