CS258: Information Theory

Fan Cheng

Spring, 2018. chengfan@sjtu.edu.cn

The normal maximizes the entropy for a given variance

If $Var(X) = a$ is fixed, then

$$h(x) \leq \frac{1}{2} \log 2\pi e a$$

The normal maximizes the entropy for a given variance

If $Var(X) = a$ is fixed, then

$$h(x) \leq \frac{1}{2} \log 2\pi e a$$

Proof.

$D(f||g) \geq 0$, where $g \sim \mathcal{N}(0, a)$

2. Entropy of a disjoint mixture

Let X_1 and X_2 be discrete random variables drawn according to probability mass functions $p_1(\cdot)$ and $p_2(\cdot)$ over the respective alphabets $\mathscr{X}_1 = \{1, 2, ..., m\}$ and $\mathscr{X}_2 = \{m+1, ..., n\}$. Let

$$X = \begin{cases} X_1, & \text{with probability } \alpha, \\ X_2, & \text{with probability } 1-\alpha. \end{cases}$$

- Find $H(X)$ in terms of $H(X_1)$, $H(X_2)$, and α.

2. Entropy of a disjoint mixture

Let X_1 and X_2 be discrete random variables drawn according to probability mass functions $p_1(\cdot)$ and $p_2(\cdot)$ over the respective alphabets $\mathscr{X}_1 = \{1, 2, ..., m\}$ and $\mathscr{X}_2 = \{m+1, ..., n\}$. Let

$$X = \begin{cases} X_1, & \text{with probability } \alpha, \\ X_2, & \text{with probability } 1-\alpha. \end{cases}$$

- Find $H(X)$ in terms of $H(X_1)$, $H(X_2)$, and α.

By definition.
Question: what if $\mathscr{X}_1 = \mathscr{X}_2$?

3. Entropy of a sum

Let X and Y be random variables that take on values $x_1, x_2, ..., x_r$ and $y_1, y_2, ..., y_s$, respectively. Let $Z = X + Y$.

(a) Show that $H(Z|X) = H(Y|X)$. Argue that if X, Y are independent, then $H(Y) \leq H(Z)$ and $H(X) \leq H(Z)$. Thus, the addition of independent random variables adds uncertainty.

(b) Give an example of (necessarily dependent) random variables in which $H(X) > H(Z)$ and $H(Y) > H(Z)$.

(c) Under what conditions does $H(Z) = H(X) + H(Y)$?

4. Entropy and pairwise independence

Let X, Y, Z be three binary $Bernoulli(\frac{1}{2})$ random variables that are pairwise independent; that is, $I(X;Y) = I(X;Z) = I(Y;Z) = 0$.

(a) Under this constraint, what is the minimum value for $H(X,Y,Z)$?

(b) Give an example achieving this minimum.

4. Entropy and pairwise independence

Let X, Y, Z be three binary $Bernoulli(\frac{1}{2})$ random variables that are pairwise independent; that is, $I(X;Y) = I(X;Z) = I(Y;Z) = 0$.

(a) Under this constraint, what is the minimum value for $H(X,Y,Z)$?

(b) Give an example achieving this minimum.

$Z = X + Y (mod 2)$

5. Subset inequality

Prove that

$$\frac{1}{2}[H(X_1, X_2) + H(X_2, X_3) + H(X_3, X_1)] \geq H(X_1, X_2, X_3)$$

5. Subset inequality

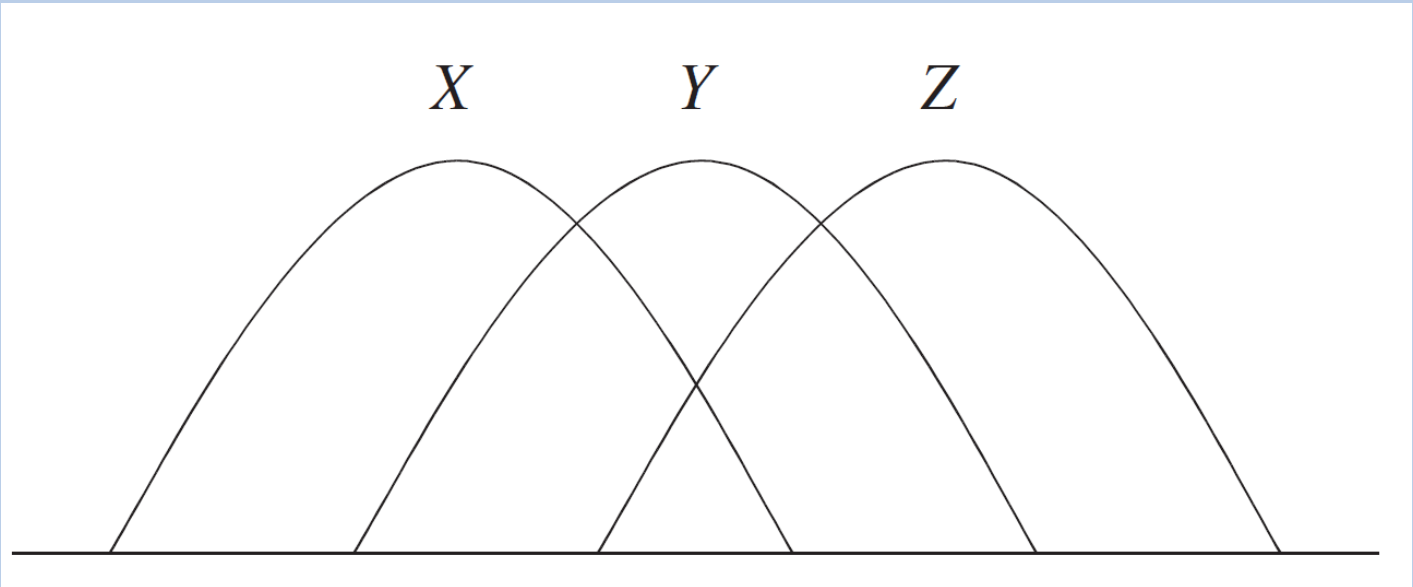
Prove that

$$\frac{1}{2}[H(X_1, X_2) + H(X_2, X_3) + H(X_3, X_1)] \geq H(X_1, X_2, X_3)$$

Information diagram

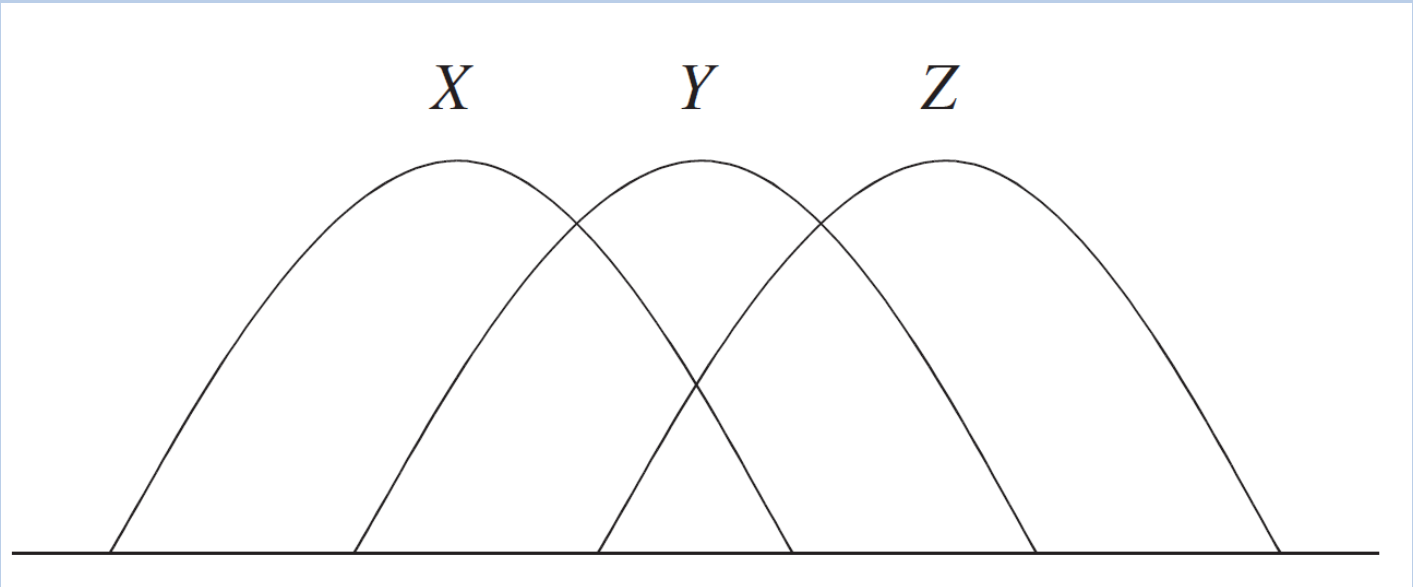
6. Information diagram for Markov chain

Verify that if $X \to Y \to Z$, then their information diagram can be simplified as



6. Information diagram for Markov chain

Verify that if $X \to Y \to Z$, then their information diagram can be simplified as



The results can be extended to $X_1 \to X_2 \to ... \to X_n$

7. Implication and Markov chain

(a) Prove that under the constraint that $X \to Y \to Z$ forms a Markov chain, $X \perp Y|Z$ and $X \perp Z$ imply $X \perp Y$.

(b) Prove that the implication in (a) continues to be valid without the Markov chain constraint.

(c) Prove that $Y \perp Z|T$ implies $Y \perp Z|(X, T)$ conditioning on $X \to Y \to Z \to T$.

8. Markov chain with 4 random variables

Let $X \to Y \to Z \to T$ form a Markov chain. Determine which of the following inequalities always hold:

(i) $I(X;T) + I(Y;Z) \geq I(X;Z) + I(Y;T)$

(ii) $I(X;T) + I(Y;Z) \geq I(X;Y) + I(Z;T)$

(iii) $I(X;Y) + I(Z;T) \geq I(X;Z) + I(Y;T)$

9. Imperfect secrecy theorem

Let X be the plain text, Y be the cipher text, and Z be the key in a secret key cryptosystem. Since X can be recovered from Y and Z, we have

$$H(X|Y,Z) = 0$$

We will show that this constraint implies

$$I(X;Y) \geq H(X) - H(Z)$$