

Problem 1

Let \mathcal{X} be a finite set and p_X be a probability distribution or a probability mass function (PMF) on \mathcal{X} . The Shannon entropy of p_X is defined as

$$H(p_X) \triangleq - \sum_{x \in \mathcal{X}} p_X(x) \log p_X(x)$$

1. Prove $\log x \leq x - 1$ and $\log \frac{1}{x} \geq 1 - x$ for all $x > 0$.
2. $\sum_{x \in \mathcal{X}} p_X(x) \log \frac{1}{p_X(x)} \leq \log |\mathcal{X}|$
3. $H(X) + H(Y) \geq H(X, Y)$ where $H(X, Y) = H(p_{X,Y})$ is the entropy of a joint PMF, $H(X) = H(p_X)$ where p_X is marginal of $p_{X,Y}$

Solution:

1. We have $\log x = \int_1^x \frac{1}{t} dt$ and $x - 1 = \int_1^x dt$. Now for $x \geq 1$ for all $t \geq 1$ we have $1 \geq \frac{1}{t}$. Hence

$$\int_1^x \frac{1}{t} dt \leq \int_1^x dt \iff \log x \leq x - 1$$

For $0 < x < 1$ we have $t < 1$ hence $\frac{1}{t} \geq 1$. Hence

$$\int_x^1 \frac{1}{t} dt \geq \int_x^1 dt \iff -\log x \geq 1 - x \iff x - 1 \geq \log x$$

Therefore $\forall x > 0$ we have $\log x \leq x - 1$.

Now we have $\log x \leq x - 1 \iff 1 - x \leq -\log x \iff 1 - x \leq \log \frac{1}{x}$.

2.

$$\begin{aligned} \sum_{x \in \mathcal{X}} p_X(x) \log \frac{1}{p_X(x)} - \log |\mathcal{X}| &= \sum_{x \in \mathcal{X}} p_X(x) \log \frac{1}{p_X(x)} - \sum_{x \in \mathcal{X}} p_X(x) \log |\mathcal{X}| \\ &= \sum_{x \in \mathcal{X}} p_X(x) \log \frac{1}{|\mathcal{X}| p_X(x)} \\ &\leq \sum_{x \in \mathcal{X}} p_X(x) \left[\frac{1}{|\mathcal{X}| p_X(x)} - 1 \right] && \text{[Using Part (1)]} \\ &= \sum_{x \in \mathcal{X}} \left[\frac{1}{|\mathcal{X}|} - p_X(x) \right] = 1 - 1 = 0 \end{aligned}$$

Hence we get

$$\sum_{x \in \mathcal{X}} p_X(x) \log \frac{1}{p_X(x)} - \log |\mathcal{X}| \iff \sum_{x \in \mathcal{X}} p_X(x) \log \frac{1}{p_X(x)} \leq \log |\mathcal{X}|$$

3.

$$\begin{aligned}
H(X) + H(Y) - H(X, Y) &= - \sum_{x \in \mathcal{X}} p_X(x) \log p_X(x) - \sum_{y \in \mathcal{Y}} p_Y(y) \log p_Y(y) \\
&\quad + \sum_{(x, y) \in \mathcal{X} \times \mathcal{Y}} p_{XY}(x, y) \log p_{XY}(x, y) \\
&= - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p_{XY}(x, y) \log p_X(x) - \sum_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} p_{XY}(x, y) \log p_Y(y) \\
&\quad + \sum_{(x, y) \in \mathcal{X} \times \mathcal{Y}} p_{XY}(x, y) \log p_{XY}(x, y) \\
&= - \sum_{(x, y) \in \mathcal{X} \times \mathcal{Y}} p_{XY}(x, y) \log \frac{p_X(x) p_Y(y)}{p_{XY}(x, y)} \\
&= \sum_{(x, y) \in \mathcal{X} \times \mathcal{Y}} p_{XY}(x, y) \log \frac{p_{XY}(x, y)}{p_X(x) p_Y(y)} \\
&\geq \sum_{(x, y) \in \mathcal{X} \times \mathcal{Y}} p_{XY}(x, y) \left[1 - \frac{p_X(x) p_Y(y)}{p_{XY}(x, y)} \right] \quad \text{[Using Part (1)]} \\
&= \sum_{(x, y) \in \mathcal{X} \times \mathcal{Y}} p_{XY}(x, y) - \sum_{(x, y) \in \mathcal{X} \times \mathcal{Y}} p_{XY}(x, y) \frac{p_X(x) p_Y(y)}{p_{XY}(x, y)} \\
&= 1 - \sum_{x \in \mathcal{X}} p_X(x) \left[\sum_{y \in \mathcal{Y}} p_Y(y) \right] \\
&= 1 - \sum_{x \in \mathcal{X}} p_X(x) = 1 - 1 = 0
\end{aligned}$$

Hence we got $H(X) + H(Y) \geq H(X, Y)$.

□

Problem 2

Let $p_X(x)$ be a PMF on \mathcal{X} . For $n \in \mathbb{N}$, $\delta > 0$, let

$$T_\delta^n(p_X) \triangleq \left\{ x^n \in \mathcal{X}^n \mid \left| \frac{N(a|x^n)}{n} - p_X(a) \right| \leq \frac{\delta p_X(a)}{\log |\mathcal{X}|} \forall a \in \mathcal{X} \right\}$$

where $N(a|x^n) = \sum_{i=1}^n \mathbb{1}_{\{x_i=a\}}$ denotes the number of occurrences of a in the sequences $x_1 x_2 \cdots x_n$.

1. Prove that

$$\sum_{x^n \notin T_\delta^n(p_X)} \prod_{i=1}^n p_X(x_i) \leq \exp \left[-\frac{2n\delta^2 \eta_{p_X}^2}{(\log |\mathcal{X}|)^2} \right]$$

where $\eta_{p_X} = \min_{a \in \mathcal{X}} \{p_X(a) \mid 0 < p_X(a) < 1\}$

2. Prove that

$$\left[1 - \exp \left(-\frac{2n\delta^2 \eta_{p_X}^2}{(\log |\mathcal{X}|)^2} \right) \right] \exp[n(H(p_X) - \delta)] \leq |T_\delta^n(p_X)| \leq \exp[n(H(p_X) + \delta)]$$

3. Prove that

$$x^n \in T_\delta^n(p_X) \implies \exp[-n(H(p_X) + \delta)] \leq \prod_{i=1}^n p_X(x_i) \leq \exp[-n(H(p_X) - \delta)]$$

Solution:

1. $\sum_{x^n \notin T_\delta^n(p_X)} \prod_{i=1}^n p_X(x_i) = \sum_{x^n \notin T_\delta^n(p_X)} p_X^n(x^n) = \Pr[x^n \notin T_\delta^n(p_X)]$. If $x^n \notin T_\delta^n(p_X)$ then there exists $a \in \mathcal{X}$ such that $\left| \frac{N(a|x^n)}{n} - p_X(a) \right| \geq \frac{\delta p_X(a)}{\log |\mathcal{X}|}$. Now $N(a|x^n) = \sum_{i=1}^n \mathbb{1}_{x_i=a}$. Hence take the indicator random variables $\mathbb{1}_{x_i=a}$ for a then $\mathbb{E}[\mathbb{1}_{x_i=a}] = p_X(a)$. Then by Hoeffding Inequality we get

$$\Pr \left[\left| \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{x_i=a} - p_X(a) \right| > \frac{\delta p_X(a)}{\log |\mathcal{X}|} \right] \leq 2 \exp \left[-2n \left(\frac{\delta p_X(a)}{\log |\mathcal{X}|} \right)^2 \right] \leq 2 \exp \left[-2n \left(\frac{\delta \eta_{p_X}}{\log |\mathcal{X}|} \right)^2 \right]$$

So

$$\begin{aligned} \Pr[x^n \notin T_\delta^n(p_X)] &\leq \sum_{a \in \mathcal{X}} \Pr \left[\left| \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{x_i=a} - p_X(a) \right| > \frac{\delta p_X(a)}{\log |\mathcal{X}|} \right] \\ &\leq \sum_{a \in \mathcal{X}} 2 \exp \left[-2n \left(\frac{\delta p_X(a)}{\log |\mathcal{X}|} \right)^2 \right] \leq 2 \exp \left[-2n \left(\frac{\delta \eta_{p_X}}{\log |\mathcal{X}|} \right)^2 \right] \\ &= 2|\mathcal{X}| \exp \left[-\frac{2n\delta^2 \eta_{p_X}^2}{\log^2 |\mathcal{X}|} \right] \end{aligned}$$

2. Using part (3) of we have In the previous part we have

$$1 \geq \sum_{x^n \in T_\delta^n(p_X)} p_X^n(x^n) \geq \sum_{x^n \in T_\delta^n(p_X)} \exp[-n(H(p_X) + \delta)] \geq |T_\delta^n(p_X)| \exp[-n(H(p_X) + \delta)]$$

Therefore we obtain

$$|T_\delta^n(p_X)| \leq \exp[n(H(p_X) + \delta)]$$

Now

$$Pr[x^n \notin T_\delta^n(p_X)] \leq 2|\mathcal{X}| \exp \left[-\frac{2n\delta^2\eta_{p_X}^2}{\log^2 |\mathcal{X}|} \right] \implies Pr[x^n \in T_\delta] \geq 1 - 2|\mathcal{X}| \exp \left[-\frac{2n\delta^2\eta_{p_X}^2}{\log^2 |\mathcal{X}|} \right]$$

And again using part (3)

$$Pr[x^n \in T_\delta] = \sum_{x^n \in T_\delta^n(p_X)} p_X^n(x^n) \leq \sum_{x^n \in T_\delta^n(p_X)} \exp[-n(H(p_X) - \delta)] \leq |T_\delta^n(p_X)| \exp[-n(H(p_X) - \delta)]$$

Therefore we have

$$|T_\delta^n(p_X)| \geq \left[1 - 2|\mathcal{X}| \exp \left(-\frac{2n\delta^2\eta_{p_X}^2}{\log^2 |\mathcal{X}|} \right) \right] \exp[n(H(p_X) - \delta)]$$

Hence we finally obtain

$$\left[1 - 2|\mathcal{X}| \exp \left(\frac{2n\delta^2\eta_{p_X}^2}{(\log |\mathcal{X}|)^2} \right) \right] \exp[n(H(p_X) - \delta)] \leq |T_\delta^n(p_X)| \leq \exp[n(H(p_X) + \delta)]$$

3. $p_X(x^n) = \prod_{i=1}^n p_X(x_i) = \prod_{a \in \mathcal{X}} p_X(a)^{N(a|x^n)}$. Now from the definition we get for all $a \in \mathcal{X}$ if $x^n \in T_\delta^n(p_X)$

$$-\frac{\delta p_X(a)}{\log |\mathcal{X}|} \leq \frac{N(a|x^n)}{n} - p_X(a) \leq \frac{\delta p_X(a)}{\log |\mathcal{X}|} \implies np_X(a) \left[1 - \frac{\delta}{\log |\mathcal{X}|} \right] \leq N(a|x^n) \leq np_X(a) \left[1 + \frac{\delta}{\log |\mathcal{X}|} \right]$$

Now we get

$$\begin{aligned} \prod_{a \in \mathcal{X}} p_X(a)^{N(a|x^n)} &\leq \prod_{a \in \mathcal{X}} p_X(a)^{np_X(a) \left[1 - \frac{\delta}{\log |\mathcal{X}|} \right]} \\ &= \prod_{x \in \mathcal{X}} \exp[np_X(a) \left[1 + \frac{\delta}{\log |\mathcal{X}|} \right] \log p_X(a)] \\ &= \exp \left[\sum_{x \in \mathcal{X}} np_X(a) \left(1 + \frac{\delta}{\log |\mathcal{X}|} \right) \log p_X(a) \right] \\ &= \exp \left[n \left(1 + \frac{\delta}{\log |\mathcal{X}|} \right) \sum_{x \in \mathcal{X}} p_X(a) \log p_X(a) \right] \\ &= \exp \left[-n \left(1 + \frac{\delta}{\log |\mathcal{X}|} \right) H(p_X) \right] \end{aligned}$$

Similarly we get

$$\prod_{a \in \mathcal{X}} p_X(a)^{N(a|x^n)} \geq \exp \left[-n \left(1 - \frac{\delta}{\log |\mathcal{X}|} \right) H(p_X) \right]$$

By Problem 1.(2) we have $H(p_X) \leq \log |\mathcal{X}|$. Hence

$$\begin{aligned} -n \left(H(p_X) + \frac{\delta H(p_X)}{\log |\mathcal{X}|} \right) &\geq -n(H(p_X) + \delta) \\ -n \left(H(p_X) - \frac{\delta H(p_X)}{\log |\mathcal{X}|} \right) &\leq -n(H(p_X) - \delta) \end{aligned}$$

Therefore we get

$$\exp[-n(H(p_X) + \delta)] \leq \prod_{i=1}^n p_X(x_i) \leq \exp[-n(H(p_X) - \delta)]$$

□

Definitions: Let $p_{X,Y}$ be a joint PMF on $\mathcal{X} \times \mathcal{Y}$ where \mathcal{X}, \mathcal{Y} are finite sets. (Essentially $p_{X,Y}(x,y) \geq 0$ and $\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p_{X,Y}(x,y) = 1$). We define the marginal of $p_{X,Y}$ on X as $p_X(x) \triangleq \sum_{y \in \mathcal{Y}} p_{X,Y}(x,y)$ for $x \in \mathcal{X}$ and marginal of $p_{X,Y}$ on Y as $p_Y(y) \triangleq \sum_{x \in \mathcal{X}} p_{X,Y}(x,y)$ for $y \in \mathcal{Y}$. or a pair $(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n$ of sequences we define $N(a,b | x^n, y^n) = \sum_{i=1}^n \mathbb{1}_{\{(x_i, y_i) = (a,b)\}}$ as the number of occurrences of (a,b) in (x^n, y^n) .

Next the joint typical set wrt $p_{X,Y}$ is defined as

$$T_\delta^n(p_{X,Y}) \triangleq \left\{ (x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n \mid \left| \frac{N(a,b | x^n, y^n)}{n} - p_{X,Y}(a,b) \right| \leq \frac{\delta p_{X,Y}(a,b)}{\log |\mathcal{X}| |\mathcal{Y}|} \forall (a,b) \in \mathcal{X} \times \mathcal{Y} \right\}$$

Problem 3

1. Prove that if $p_{X,Y}(a,b) = 0$ for some $(a,b) \in \mathcal{X} \times \mathcal{Y}$ and $(x^n, y^n) \in T_\delta^n(p_{X,Y})$ then $N(a,b | x^n, y^n) = 0$. In other words, a pair that - probability does not occur in any typical pair of sequences.
2. Let $\eta_{p_{X,Y}} = \min_{(x,y) \in \mathcal{X} \times \mathcal{Y}} \{p_{X,Y}(x,y) \mid 0 < p_{X,Y}(x,y) < 1\}$. Use the Hoeffding Inequality to prove that

$$\sum_{(x^n, y^n) \notin T_\delta^n(p_{X,Y})} p_{X,Y}^n(x^n, y^n) \leq 2|\mathcal{X}| |\mathcal{Y}| \exp \left[-\frac{2n\delta^2 \eta_{p_{X,Y}}^2}{(\log |\mathcal{X}| |\mathcal{Y}|)^2} \right]$$

Hoeffding Inequality: Let Z_1, \dots, Z_m are independent and identically distributed random variables for which $P[a \leq Z_i \leq b] = 1$ for ever $1 \leq i \leq m$ and $\mu = \mathbb{E}[Z_i]$. Then for every $\epsilon > 0$

$$P \left[\left| \frac{1}{m} \sum_{i=1}^m Z_i - \mu \right| > \epsilon \right] \leq 2 \exp \left[-2m \frac{\epsilon^2}{(b-a)^2} \right]$$

3. For any $(x^n, y^n) \in T_\delta^n(p_{X,Y})$ prove that

$$2^{-n[H(p_{X,Y}) + \delta]} \leq p_{X,Y}^n(x^n, y^n) = \prod_{i=1}^n p_{X,Y}(x_i, y_i) \leq 2^{-n[H(p_{X,Y}) - \delta]}$$

4. Prove that

$$(1 - \tilde{\delta}) 2^{n[H(p_{X,Y}) - \delta]} \leq |T_\delta^n(p_{X,Y})| \leq 2^{n[H(p_{X,Y}) + \delta]}$$

$$\text{where } \tilde{\delta} = 2|\mathcal{X}| |\mathcal{Y}| \exp \left[-\frac{2n\delta^2 \eta_{p_{X,Y}}^2}{(\log |\mathcal{X}| |\mathcal{Y}|)^2} \right]$$

5. Prove that $(x^n, y^n) \in T_\delta^n(p_{X,Y})$ then $x^n \in T_\delta^n(p_X)$ and $y^n \in T_\delta^n(p_Y)$.

Solution:

□

Definitions: Suppose p_{XY} is a probability distribution (probability mass function (PMF)) on $\mathcal{X} \times \mathcal{Y}$. We recall the condition distribution $p_{Y|X}(y|x) = \frac{p_{XY}(x,y)}{p_X(x)}$ and for a pair $(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n$ of sequence $(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n$ of sequences $p_{Y|X}^n(y^n|x^n) = \prod_{i=1}^n p_{Y|X}(y_i|x_i)$

We define

$$H(Y|X = x) \triangleq H(p_{Y|X}|X = x) = - \sum_{y \in \mathcal{Y}} p_{Y|X}(y|x) \log p_{Y|X}(y|x)$$

and

$$H(Y|X) = H(p_{Y|X}|p_X) \triangleq \sum_{x \in \mathcal{X}} p_X(x) h(Y|X = x)$$

For any $x^n \in \mathcal{X}^n$ define the conditional typical set of x^n as

$$T_\delta^n(p_{Y|X}) = \{y^n \in \mathcal{Y}^n \mid (x^n, y^n) \in T_\delta^n(p_{XY})\}$$

Problem 4

1. Prove that $\sum_{y \in \mathcal{Y}} p_{Y|X}(y|x) = 1$
2. Prove that $H(Y|X) = H(X, Y) - H(X)$ and $H(Y|X) \geq 0$
3. Prove that Verify that if $x^n \notin T_\delta^n(p_X)$ then $T_\delta^n(p_{XY}|x^n) = \emptyset$
4. Suppose $x^n \in T_\delta^n(p_X)$ and $y^n \in T_\delta^n(p_{XY}|x^n)$ prove that

$$2^{-n[H(Y|X)+2\delta]} \leq p_{Y|X}^n(y^n|x^n) \leq 2^{-n[H(Y|X)-2\delta]}$$

5. Prove that if $x^n \in T_\delta^n(p_X)$ then

$$\sum_{y^n \in T_{2\delta}^n(p_{XY}|x^n)} p_{Y|X}^n(y^n|x^n) \geq 1 - 2|\mathcal{X}||\mathcal{Y}| \exp \left[-\frac{2n\delta^2}{(\log |\mathcal{X}||\mathcal{Y}|)^2} \eta_{p_{Y|X}} \right]$$

where $\eta_{p_{Y|X}} = \min_{(x,y) \in \mathcal{X} \times \mathcal{Y}} \{p_{Y|X}(y|x) \mid 0 < p_{Y|X}(y|x) < 1\}$

6. Suppose $x^n \in T_\delta^n(p_X)$ then

$$(1 - \tilde{\delta})2^{n[H(Y|X)-4\delta]} \leq |T_\delta^n(p_{XY}|x^n)| \leq 2^{n[H(Y|X)+4\delta]}$$

where $\tilde{\delta} = 2|\mathcal{X}||\mathcal{Y}| \exp \left[-\frac{2n\delta^2}{(\log |\mathcal{X}||\mathcal{Y}|)^2} \eta_{p_{Y|X}} \right]$

Solution:

□