ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

School of Computer and Communication Sciences

**Handout 34** | Information Theory and Coding

Solution to extra homework | Jan. 06, 2020

---

Problem 1. In the lecture, the professor frequently uses the following lemma :

Lemma 1. For any PMF $Q_{U,V}$ defined on finite alphabet $\mathcal{U}$ and $\mathcal{V}$, if $u^n \in \mathcal{T}(Q_U, n, \delta)$ and $V^n \sim Q_V^n$ (i.e. $V^n$ is sampled i.i.d from the marginal distribution $Q_V$ independent of $u^n$) then for large enough $n$ we have

$$\mathbb{P}\left((u^n, V^n) \in \mathcal{T}(Q_{U,V}, n, 3\delta)\right) \geq 2^{-n(I(V;U)-3\delta-|\mathcal{U}|\delta-|\mathcal{U}||\mathcal{V}|\delta)}.$$

Several students ask us about the proof of this lemma, as the professor left it as an exercise. In this problem, we will prove this lemma.

For each $a \in \mathcal{U}$, we define a set of indices $N_{u^n}(a) = \{i \mid u_i = a, i \in [1, n]\}$, i.e. the indices where $u_i = a$. Using this definition, for each $u^n$, we can define the conditional typical set as

$$\mathcal{T}(Q_{V|U}, u^n, \delta) = \left\{ v^n \,\middle|\, \forall_{a \in \mathcal{U}, b \in \mathcal{V}} \left| \frac{\sum_{i \in N_{u^n}(a)} \mathbb{1}\{v_i = b\}}{|N_{u^n}(a)|} - Q_{V|U}(b|a) \right| \leq \delta \right\}.$$

Informally, this definition says that if we group the $V^n$ according to the value of $u^n$ in the corresponding position, then its empirical distribution is close to the actual distribution.

a) Show that if $u^n \in \mathcal{T}(Q_U, n, \delta)$ and $v^n \in \mathcal{T}(Q_{V|U}, u^n, \delta)$ then (1) $(u_n, v_n) \in \mathcal{T}(Q_{U,V}, n, 3\delta)$ and (2) $v^n \in \mathcal{T}(Q_V, n, |\mathcal{U}|\delta + \delta)$.

*First of all, it is easy to see that from the definition of $N_{u^n}(a)$ that*

$$\sum_{i \in [1,n]} \mathbb{1}\{u_i = a, v_i = b\} = \sum_{i \in N_{u^n}(a)} \mathbb{1}\{v_i = b\}.$$

*For every $a \in \mathcal{U}$ and $b \in \mathcal{V}$, consider the following inequality*

$$\begin{aligned}
\left| \frac{\sum_{i \in [1,n]} \mathbb{1}\{u_i = a, v_i = b\}}{n} - Q_{U,V}(a,b) \right| &\leq \left| \frac{\sum_{i \in N_{u^n}(a)} \mathbb{1}\{v_i = b\}}{n} - Q_{V|U}(b|a)\frac{|N_{u^n}(a)|}{n} \right| \\
&\quad + \left| Q_{V|U}(b|a)\frac{|N_{u^n}(a)|}{n} - Q_{U,V}(a,b) \right| \\
&= \frac{|N_{u^n}(a)|}{n} \left| \frac{\sum_{i \in N_{u^n}(a)} \mathbb{1}\{v_i = b\}}{|N_{u^n}(a)|} - Q_{V|U}(b|a) \right| \\
&\quad + Q_{V|U}(b|a) \left| \frac{|N_{u^n}(a)|}{n} - Q_U(a) \right| \\
&\leq 2\delta
\end{aligned}$$

*where the first inequality is due to the triangle inequality, and the second inequality is due to our assumption on $u^n$ and $v^n$. This proves statement (1).*

*Now, for every $b \in \mathcal{V}$ consider the following inequality*

$$\begin{aligned}
\left|\frac{\sum_{i\in[1,n]} \mathbb{1}\{v_i = b\}}{n} - Q_V(b)\right| &= \left|\sum_{a\in\mathcal{U}} \frac{|N_{u^n}(a)|}{n} \frac{\sum_{i\in N_{u^n}(a)} \mathbb{1}\{v_i = b\}}{|N_{u^n}(a)|} - \sum_{a\in\mathcal{U}} \frac{|N_{u^n}(a)|}{n} Q_{V|U}(b|a)\right.\\
&\quad \left. + \sum_{a\in\mathcal{U}} \frac{|N_{u^n}(a)|}{n} Q_{V|U}(b|a) - \sum_{a\in\mathcal{U}} Q_U(a) Q_{V|U}(b|a)\right|\\
&\le \sum_{a\in\mathcal{U}} \frac{|N_{u^n}(a)|}{n} \left|\frac{\sum_{i\in N_{u^n}(a)} \mathbb{1}\{v_i = b\}}{|N_{u^n}(a)|} - Q_{V|U}(b|a)\right|\\
&\quad + \sum_{a\in\mathcal{U}} Q_{V|U}(b|a) \left|\frac{|N_{u^n}(a)|}{n} - Q_U(a)\right|\\
&\le \delta + |\mathcal{U}|\delta
\end{aligned}$$

*where the first inequality is the triangle inequality and the second inequality is due to the assumptions. This implies statement (2).*

b) In the class, we have showed the proof of the lower bound on the size of typical set by using the law of large number. We can get an analogous result using the Hoeffding's bound. Start by showing that for if $V^{|N_{u^n}(a)|} \sim Q_{V|U=a}^{|N_{u^n}(a)|}$ then

$$1 - 2|\mathcal{V}| \exp\left(-2\delta^2 |N_{u^n}(a)|\right) \le \mathbb{P}\left(V^{|N_{u^n}(a)|} \in \mathcal{T}(Q_{V|U=a}, |N_{u^n}(a)|, \delta)\right)$$

Argue why the lower bound implies that for large enough $|N_{u^n}(a)|$ we have:

$$|N_{u^n}(a)| \left(H(V|X = a) - 2\delta\right) \le \log_2\left(|\mathcal{T}(Q_{V|U=a}, |N_{u^n}(a)|, \delta)|\right)$$

Note that the random variable $V^{|N_{u^n}(a)|}$ does not have anything to do with the $V^n$ in the lemma as it has different distribution. However, the set $\mathcal{T}(Q_{V|U=a}, |N_{u^n}(a)|, \delta)$ is a combinatoric object, and it sizes does not depend on the underlying distribution.

*From the definition of typical set, we have :*

$$\begin{aligned}
\mathbb{P}(V^{|N_{u^n}(a)|} \notin \mathcal{T}(Q_{V|U=a}, |N_{u^n}(a)|, \delta)) &= \mathbb{P}\left(\bigcup_{b\in\mathcal{V}} \left|\frac{\sum_{i\in[1,|N_{u^n}(a)|]} \mathbb{1}\{V_i = b\}}{|N_{u^n}(a)|} - Q_{V|U=a}(b)\right| > \delta\right)\\
&\le \sum_{b\in\mathcal{V}} \mathbb{P}\left(\left|\frac{\sum_{i\in[1,|N_{u^n}(a)|]} \mathbb{1}\{V_i = b\}}{|N_{u^n}(a)|} - Q_{V|U=a}(b)\right| > \delta\right)\\
&\le \sum_{b\in\mathcal{V}} 2\exp\left(-2\delta^2 |N_{u^n}(a)|\right)\\
&= 2|\mathcal{V}|2\exp\left(-2\delta^2 |N_{u^n}(a)|\right)
\end{aligned}$$

*where the first inequality is due to union bound, and the second inequality is due to Hoeffding Bound. This implies the required form in the problem statement.*

*We have :*

$$\begin{aligned}
1 - 2|\mathcal{V}| \exp\left(-2\delta^2 |N_{u^n}(a)|\right) &\le \mathbb{P}\left(V^{|N_{u^n}(a)|} \in \mathcal{T}(Q_{V|U=a}, |N_{u^n}(a)|, \delta)\right)\\
&\le \sum_{v^{|N_{u^n}(a)|}\in\mathcal{T}(Q_{V|U=a}, |N_{u^n}(a)|, \delta)} \mathbb{P}(v^{|N_{u^n}(a)|})\\
&\le \sum_{v^{|N_{u^n}(a)|}\in\mathcal{T}(Q_{V|U=a}, |N_{u^n}(a)|, \delta)} 2^{-|N_{u^n}(a)|(H(V|U=a)-\delta)}\\
&= 2^{-|N_{u^n}(a)|(H(V|U=a)-\delta)} |\mathcal{T}(Q_{V|U=a}, |N_{u^n}(a)|, \delta)|
\end{aligned}$$

*Thus we only need to check that for large enough $|N_{u^n}(a)|$ we have:*

$$\frac{\log_2\left(1 - 2\,|\mathcal{V}|\exp\left(-2\delta^2|N_{u^n}(a)|\right)\right)}{|N_{u^n}(a)|} \leq \delta.$$

*This is possible because the numerator of the LHS goes to* $0$*, but the denumerator goes to* $\infty$*.*

c) Show that if $u^n \in \mathcal{T}\left(Q_U, n, \delta\right)$ then for large enough $n$ we have:

$$2^{n(H(V|U)-2\delta-|\mathcal{U}||\mathcal{V}|\delta)} \leq \left|\mathcal{T}(Q_{V|U}, u^n, \delta)\right|.$$

[Hint : $\left|\mathcal{T}(Q_{V|U}, u^n, \delta)\right| = \prod_{a\in\mathcal{U}}\left|\mathcal{T}(Q_{V|U=a}, |N_{u^n}(a)|, \delta)\right|$]

*Accepting the hint and using the result of (b), we have*

$$\begin{aligned}\log_2\left|\mathcal{T}(Q_{V|U}, u^n, \delta)\right| &\geq \sum_{a\in\mathcal{U}} N_{u^n}(a)H(V|U=a) - \sum_{a\in\mathcal{U}} N_{u^n}(a)2\delta \\ &\geq \sum_{a\in\mathcal{U}} nQ_U(a)H(V|U=a) - \sum_{a\in\mathcal{U}} n\delta H(V|U=a) - n2\delta \\ &\geq nH(V|U) - n|\mathcal{V}||\mathcal{U}|\delta - n2\delta\end{aligned}$$

d) Show that if $u^n \in \mathcal{T}\left(Q_U, n, \delta\right)$ then for large enough $n$

$$2^{-n(I(V;U)-3\delta-|\mathcal{U}|\delta-|\mathcal{U}||\mathcal{V}|\delta)} \leq \mathbb{P}\left(V^n \in \mathcal{T}(Q_{V|U}, u^n, \delta)\right)$$

Argue why this implies the statement of the lemma.

*From (a), we know that if* $u^n \in \mathcal{T}\left(Q_U, n, \delta\right)$ *and* $v^n \in \mathcal{T}(Q_{V|U}, u^n, \delta)$ *then* $v^n \in \mathcal{T}(Q_V, n, |\mathcal{U}|\delta + \delta)$*. This observation will be used in the following inequality*

$$\begin{aligned}\mathbb{P}\left(V^n \in \mathcal{T}(Q_{V|U}, u^n, \delta)\right) &= \sum_{v^n\in\mathcal{T}(Q_{V|U}, u^n, \delta)} \mathbb{P}(v^n) \\ &\geq 2^{-nH(V)-n\delta-n|\mathcal{U}|\delta}\left|\mathcal{T}(Q_{V|U}, u^n, \delta)\right| \\ &\geq 2^{-nH(V)+nH(V|U)-3n\delta-n\delta|\mathcal{U}|-n\delta\mathcal{U}|\mathcal{V}|} \\ &\geq 2^{-n(I(V;U)-3\delta-\delta|\mathcal{U}|-\delta\mathcal{U}|\mathcal{V}|)}\end{aligned}$$

*This proves the inequality. By the result of (a), we also know that*

$$\mathbb{P}\left((u_n, V^n) \in (u_n, V_n) \in \mathcal{T}(Q_{U,V}, n, 3\delta)\right) \geq \mathbb{P}\left(V^n \in \mathcal{T}(Q_{V|U}, u^n, \delta)\right)$$

*as every* $V^n$ *which is included in the event on RHS is also included in the event in LHS.*