EE376A Information Theory Lecture 13 - 02/19/2015

Lecture 13

Lecturer: Tsachy Weissman *Scribe: Huseyin A. Inan*

1 Method of Types (Sec. 11.1 in C & T)

$x^n = (x_1, \ldots, x_n), \qquad x_i \in \mathcal{X} = \{1, 2, \ldots, r\}.$

Definition 1. *The empirical distribution or type of x^n is the vector $(P_{x^n}(1), P_{x^n}(2), \ldots P_{x^n}(r))$ of relative frequencies $P_{x^n}(a) = \dfrac{N(a|x^n)}{n}$ where $N(a|x^n) = \sum_{i=1}^n \mathbb{1}_{\{x_i = a\}}$.*

Let $\mathscr{P}_n$ denote the collection of all empirical distributions of sequences of length n.

Example 2. $\mathcal{X} = \{0, 1\}$

$$\mathscr{P}_n = \left\{ (0,1), \left(\frac{1}{n}, \frac{n-1}{n}\right), \left(\frac{2}{n}, \frac{n-2}{n}\right), \ldots, (1,0) \right\}.$$

Definition 3. *If $P \in \mathscr{P}_n$ (probabilities are integer multiples of $1/n$), the type class or type of P is $T(P) = \{x^n : P_{x^n} = P\}$. The type class of x^n is $T_{x^n} = T(P_{x^n}) = \{\tilde{x} : P_{\tilde{x}^n} = P_{x^n}\}$.*

Example 4. $\mathcal{X} = \{a, b, c\}$, $n = 5$, $x^n = (aacba)$

Then $P_{x^n} = (3/5, 1/5, 1/5)$

$T_{x^n} = \{aaabc, aaacb, \ldots, cbaaa\}$

$|T_{x^n}| = \binom{5}{3\ 1\ 1} = \dfrac{5!}{3!1!1!} = 20.$

Theorem 5. $|\mathscr{P}_n| \leq (n+1)^{r-1}$

Proof Type of x^n is determined by $(N(1|x^n), N(2|x^n), \ldots, N(r|x^n))$. Each component can assume no more than $n+1$ values $(0 \leq N(i|x^n) \leq n)$, (and the last component is dictated by the others). □

E.g. : For $\mathcal{X} = \{0, 1\}$, $\quad |\mathscr{P}_n| = n + 1 = (n+1)^{r-1}$.

1.1 Notation

- $Q = \{Q(x)\}_{x \in \mathcal{X}}$ is a PMF, write $H(Q)$ for $H(X)$ when $X \sim Q$.
- $Q^n(x^n) = \prod_{i=1}^n Q(x_i), \qquad S \subseteq \mathcal{X}^n \quad Q^n(S) = \sum_{x^n \in S} Q^n(x^n)$.

Theorem 6. $\forall x^n : Q^n(x^n) = 2^{-n[H(P_{x^n})+D(P_{x^n}||Q)]}$.

Proof

$$\begin{aligned}
Q^n(x^n) &= \prod_{i=1}^{n} Q(x_i) = 2^{\sum_{i=1}^{n} \log Q(x_i)} \\
&= 2^{\sum_{a\in\mathcal{X}} N(a|x^n)\log Q(a)} \\
&= 2^{n\sum_{a\in\mathcal{X}} \frac{N(a|x^n)}{n}\log Q(a)} \\
&= 2^{-n\sum_{a\in\mathcal{X}} \frac{N(a|x^n)}{n}\log \frac{1}{Q(a)}} \\
&= 2^{-n\left[\sum_{a\in\mathcal{X}} P_{x^n}(a)\log \frac{1}{Q(a)}\right]} \\
&= 2^{-n\left[\sum_{a\in\mathcal{X}} P_{x^n}(a)\log \frac{1}{P_{x^n}(a)} + \sum_{a\in\mathcal{X}} P_{x^n}(a)\log \frac{P_{x^n}(a)}{Q(a)}\right]} \\
&= 2^{-n[H(P_{x^n})+D(P_{x^n}||Q)]}.
\end{aligned}$$

□

Theorem 7. $\forall P \in \mathscr{P}_n$

$$\frac{1}{(n+1)^{r-1}} 2^{nH(P)} \leq |T(P)| \leq 2^{nH(P)}.$$

NOTE: $|T(P)| = \binom{n}{nP(1)\ nP(2)\ \dots\ nP(r)} = \dfrac{n!}{\prod_{a\in\mathcal{X}}(nP(a))!}$.

Proof UPPER BOUND:

$$\begin{aligned}
1 \geq P^n(T(P)) &= \sum_{x^n\in T(P)} P^n(x^n) \\
&= |T(P)| 2^{-n[H(P)+D(P||P)]} \\
&= |T(P)| 2^{-nH(P)}.
\end{aligned}$$

□

For the lower bound we will use a Lemma.

Lemma: $\forall P, Q \in \mathscr{P}_n : P^n(T(P)) \geq P^n(T(Q))$.

Proof

$$\begin{aligned}
\frac{P^n(T(P))}{P^n(T(Q))} = \frac{|T(P)|\prod_{a\in\mathcal{X}} P(a)^{[nP(a)]}}{|T(Q)|\prod_{a\in\mathcal{X}} P(a)^{[nQ(a)]}} &= \frac{\binom{n}{nP(1)\ \dots\ nP(r)}}{\binom{n}{nQ(1)\ \dots\ nQ(r)}} \prod_{a\in\mathcal{X}} P(a)^{n[P(a)-Q(a)]} \\
&= \prod_{a\in\mathcal{X}} \frac{(nQ(a))!}{(nP(a))!} P(a)^{n[P(a)-Q(a)]}
\end{aligned}$$

NOTE: $\dfrac{m!}{n!} \geq n^{m-n}$

If $m > n$, then $\frac{m!}{n!} = m(m-1)\dots(n+1) \geq n^{m-n}$.

If $n > m$, then $\frac{m!}{n!} = \frac{1}{n(n-1)\dots(m+1)} \geq \left(\frac{1}{n}\right)^{n-m} = n^{m-n}$.

Therefore,

$$\begin{aligned}\prod_{a\in\mathcal{X}} \frac{(nQ(a))!}{(nP(a))!} P(a)^{n[P(a)-Q(a)]} &\geq \prod_{a\in\mathcal{X}} (nP(a))^{n[Q(a)-P(a)]}\, P(a)^{n[P(a)-Q(a)]} \\ &= \prod_{a\in\mathcal{X}} n^{n[P(a)-Q(a)]} = n^{n\sum_{a\in\mathcal{X}}[P(a)-Q(a)]} = 1.\end{aligned}$$

□

PROOF OF LOWER BOUND:

$$\begin{aligned}1 = \sum_{Q\in\mathscr{P}_n} P^n\left(T(Q)\right) &\leq |\mathscr{P}_n| \max_Q P^n\left(T(Q)\right) \\ &= |\mathscr{P}_n| P^n\left(T(P)\right) \\ &= |\mathscr{P}_n||T(P)|2^{-n[H(P)+D(P||P)]} \\ &\leq (n+1)^{r-1}\,|T(P)|2^{-nH(P)}.\end{aligned}$$

Theorem 8. $\forall P \in \mathscr{P}_n, Q$

$$\frac{1}{(n+1)^r} 2^{-nD(P||Q)} \leq Q^n\left(T(P)\right) \leq 2^{-nD(P||Q)}.$$

Proof $Q^n\left(T(P)\right) = |T(P)|2^{-n[H(P)+D(P||Q)]}$.

Now bound $|T(P)|$ as in previous theorem. □

We will write $\alpha_n \doteq \beta_n$: "equality to first order in the exponent"

$\iff \frac{1}{n}\log\frac{\alpha_n}{\beta_n} \xrightarrow{n\to\infty} 0$

$\iff |\frac{1}{n}\log\alpha_n - \frac{1}{n}\log\beta_n| \xrightarrow{n\to\infty} 0$

E.g. : $\alpha_n \doteq 2^{nJ} \iff \alpha_n = 2^{n(J+\epsilon_n)}$ where $\epsilon_n \xrightarrow{n\to\infty} 0$.

1.2 Recap

- $|\mathscr{P}_n| \leq (n+1)^{r-1}$
- $Q^n(x^n) = 2^{-n[H(P_{x^n})+D(P_{x^n}||Q)]}$
- $|T(P)| \doteq 2^{nH(P)}, \quad P \in \mathscr{P}_n$
- $Q^n\left(T(P)\right) \doteq 2^{-nD(P||Q)}$