Statistics G 8243
Ioannis Kontoyiannis
Thursday, February 12, 2009
Handout #7

TYPES – Definitions and basic properties

The **type** $\hat{P}_{x_1^n}$ of a string $x_1^n = (x_1, x_2, \ldots, x_n)$ taking values in a finite alphabet $A = \{a_1, a_2, \ldots, a_m\}$ is its empirical distribution:

$$\hat{P}_{x_1^n}(a) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}_{\{x_i = a\}}\ , \qquad a \in A.$$

An n-**type** P is a distribution on A such that for each $a \in A$, $P(a) = k/n$ for an integer k. The collection of all n-types is denoted $\mathcal{P}_n$.

The **type class** $T(P)$ of an n-type P is the collection of all strings x_1^n with type P:

$$T(P) = \{x_1^n \in A^n \ : \ \hat{P}_{x_1^n} = P\}.$$

Properties:

1. $$|\mathcal{P}_n| \leq (n+1)^m$$

2. For every distribution Q and every $x_1^n \in A^n$:

$$Q^n(x_1^n) = 2^{-n[H(\hat{P}_{x_1^n}) + D(\hat{P}_{x_1^n} \| Q)]}$$

3. If $P \in \mathcal{P}_n$ and $x_1^n \in T(P)$:

$$P^n(x_1^n) = 2^{-nH(P)}$$

4. If $P \in \mathcal{P}_n$, then $|T(P)| \doteq 2^{nH(P)}$:

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

5. For every distribution Q and every $P \in \mathcal{P}_n$, we have $Q^n(T(P)) \doteq 2^{-nD(P\|Q)}$:

$$\frac{1}{(n+1)^m} 2^{-nD(P\|Q)} \leq Q^n(T(P)) \leq 2^{-nD(P\|Q)}$$