Information Theory Notes

Charles Zheng and Yuval Benjamini

January 28, 2016

1 Proof of Key Result

We give the proof with minimal context or motivation. See the paper for more details.

Fix integer $K \geq 2$. Let $p^{[d]}(x, y)$ be a sequence of probability density functions, where x is of dimension $p^{[d]}$ and y is of dimension $q^{[d]}$. Let $p^{[d]}(x)$ and $p^{[d]}(y)$ denote the marginal densities, and let

$$p^{[d]}(y|x) = p^{[d]}(x, y)/p^{[d]}(y).$$

Let $(X^{([d],i)}, Y^{([d],i)})$ be iid random variates from $p^{[d]}(x, y)$ for $i = 0, \ldots, K-1$; we will supress the superscripts $[d]$ and/or (i) when convenient. Recall the definitions of entropy,

$$H(X) = -\int p(x) \log p(x) dx,$$

and mutual information

$$I(X; Y) = \int p(x, y) \log \frac{p(x, y)}{p(x)p(y)} dxdy.$$

Furthermore, define the K-class average Bayes error as

$$\text{ABE}_K \Pr[p(Y^{(0)}|X^{(0)}) < \max_{i=1}^{K-1} p(Y^{(0)}|X^{(i)})].$$

Define

$$u^{[d]}(x, y) = \log p^{[d]}(x, y) - \log p^{[d]}(x) - \log p^{[d]}(y).$$