Assignment 9

Q.1 – Let C be a finite context set and let $c_1, c_2, \dots c_n \in C$ be an arbitary sequence of contexts.

1) Show that:

$$\sum_{c \in C} \sqrt{\sum_{t=1}^{n} \mathbb{I}\{c_t = c\}} \leq \sqrt{n|C|}$$

Proof: Since square root is concave function, Hence by Jensen's inequality:

$$\frac{\sum_{i=1}^{n} \sqrt{(a_i x_i)^2}}{n} \leq \sqrt{\frac{\sum_{i=1}^{n} (a_i x_i)^2}{n}}$$

$\therefore$ for $a_i = 1$

$$\sqrt{\frac{\sum_{i=1}^{n} x_i}{n}} \geq \frac{\sum_{i=1}^{n} \sqrt{x_i}}{n} \quad \text{—— (1)}$$

$x($ . .

Now,

$$\sum_{c \in C} \sqrt{\sum_{t=1}^{n} I\{c_t = c\}} = \frac{1}{|C|} \sum_{c \in C} |C| \sqrt{\sum_{t=1}^{n} I\{c_t = c\}}$$

$$\leq |C| \sqrt{\sum_{c \in C} \frac{1}{|C|} \sum_{t=1}^{n} I\{c_t = c\}} \quad \Big[\text{from eq } 1$$

$$\text{—— } 2$$

Now Since,

$$\sum_{c \in C} \sum_{t=1}^{n} I\{c_t = c\} = n$$

Inputting this in eq 2, we get:

$$\sum_{c \in C} \sqrt{\sum_{t=1}^{n} I\{c_t = c\}} \leq |C| \sqrt{\frac{n}{|C|}}$$

$$\leq \sqrt{n|C|}$$

Hence Proved

2) Assume that n is an integer multiple of |C|. Show that the choice that maximizes the right-hand side of the previous inequality is the one when each context occurs $n/|C|$ times.

Proof: Let $n = k|C|$, k is an integer

$\Rightarrow k = \dfrac{n}{|C|}$ is the number of times a context occurs.

$$k = \sum_{t=1}^{n} I\{c_t = c\} = \frac{n}{|C|}$$

Putting this in the L.H.S. of the previous inequality, we get

$$\underbrace{\sum_{c \in C}}_{|C|} \sqrt{\underbrace{\sum_{t=1}^{n} I\{c_t = c\}}_{\sqrt{\frac{n}{|C|}}}} \leq \sqrt{n|C|}$$

$$\Rightarrow k|C| = \sqrt{n|C|} \quad \text{when } k = \frac{n}{|C|}$$

Hence Proved.