1. What is $\lim_{\beta \to \infty} \omega$? $\rightarrow T = \{T_1, T_2, \dots, T_N\}$ and $\omega = \frac{1}{\beta} \log \sum_{k=1}^{N} \exp\{\beta T_k\}$ where $\beta > 0$

Compute scalar quantity $z = \sum_{k=1}^{N} T_k \frac{\exp(\beta T_k)}{\sum_{i=1}^{N} \exp(\beta T_i)}$ where $\beta > 0$

What is $\lim_{\beta \to \infty} z$?

$$\lim_{\beta \to \infty} \sum_{k=1}^{N} T_k \frac{\exp(\beta T_k)}{\sum_{i=1}^{N} \exp(\beta T_i)}$$

$$z = \frac{e^{\beta T_0}(T_1 e^{\beta(T_1 - T_0)} + \dots + T_N e^{\beta(T_N - T_0)})}{e^{\beta T_0}(e^{\beta(T_1 - T_0)} + \dots + T_N e^{\beta(T_N - T_0)})}$$

$$z = \frac{T_0 \cancel{e^{\beta(T_N - T_0)}}}{\cancel{e^{\beta(T_N - T_0)}}} = \boxed{T_0}$$

$$= \lim_{\beta \to \infty} \frac{1}{\beta} \log \sum_{k=1}^{N} \exp\{\beta T_k\}$$

$$= \frac{\log e^{\beta T_0} + \log[e^{\beta(T_1 - T_0)} + e^{\beta(T_2 - T_0)}]}{\beta}$$

$$\omega = \frac{\log e^{\beta T_0} + \log(1)}{\beta} = \frac{\beta T_0 + 0}{\beta} = \boxed{T_0}$$

Among set T

2a. For sigmoid activation function $z(\upsilon) = \sigma(\upsilon) = \frac{1}{1 + \exp(-\upsilon)}$,

and BCE loss function $\ell(y, z(\upsilon)) = y \log \frac{y}{z} + (1-y) \log \frac{1-y}{1-z}$.

$$= y(\log(y) - \log(z)) + (1-y)\log\left(\frac{1-y}{1-z}\right)$$

$$= y\log(y) - y\log(z) + \underbrace{(1-y)\log\left(\frac{1-y}{1-z}\right)}$$

$$= \quad \dots \quad + \underbrace{(1-y)(\log(1-y) - \log(1-z))}_{\text{Define as variable } \upsilon}$$

$$= y\log(y) - y\log(z) + (1-y)\upsilon$$

$$= -y\upsilon + y(\log(y) + \upsilon) + y\log(z)$$

$$= -y\upsilon + \log(1 + \exp\{\upsilon\}) + y\log(z) + \text{terms independent of } \upsilon$$

2b. https://colab.research.google.com/drive/1kUA1Q3sHpXvrXg5GFzVG38KQUy5rVAh1?usp=sharing

3. Schwarz Inequality: $\left(\sum_{k=1}^{N} |x_k y_{k_1}|\right)^2 \leq \sum_{k=1}^{N} |x_k|^2 \sum_{k=1}^{N} |y_{k_2}|^2$

For $c_n \geq 0$, $\sum_{n=1}^{N} c_n = 1$ and a vector $x \in \mathbb{R}^N$ with elements $x_n \geq 0$,

$\sum_{n=1}^{N} c_n x_n^2 \geq \left(\sum_{n=1}^{N} c_n x_n\right)^2$:

$\rightarrow$ From this, we can say that $c_n = v_1$ and $\sqrt{c_n} = v_2$

v_1 and v_2 would be worthy substitutions for y_{k_1} and y_{k_2} respectively.

$[k = n]$

$$\left(\sum_{n=1}^{N} |(x_n)(c_n)|\right)^2 \leq \sum_{n=1}^{N} |x_n|^2 \sum_{n=1}^{N} |\sqrt{c_n}|^2$$

$$\left(\sum_{n=1}^{N} |(x_n)(c_n)|\right)^2 \le \sum_{n=1}^{N} |x_n|^2 \sum_{n=1}^{N} |\sqrt{c_n}|^2$$

$$\left(\sum_{n=1}^{N} c_n x_n\right)^2 \le \sum_{n=1}^{N} x_n^2 c_n$$