Bernoulli MLE Derivation

$$\hat{\theta}_{MLE} = \arg\max_{\theta} \overline{P(D \mid \theta)}$$

$$= \arg\max_{\theta} \prod_{i=1}^{n} \theta^{x_i} (1-\theta)^{1-x_i}$$

$$= \arg\max_{\theta} \underbrace{\sum_{i=1}^{n} x_i \log\theta + (1-x_i)\log(1-\theta)}_{J(\theta)}$$

$$\rightarrow \frac{\partial J(\theta)}{\partial \theta} = \sum_{i=1}^{n} \frac{x_i}{\theta} - \frac{(1-x_i)}{1-\theta} = 0 \qquad x_i = \begin{cases} 1 & H \\ 0 & T \end{cases}$$

$$\frac{\sum_{i=1}^{n} x_i}{\theta} = \frac{\sum_{i=1}^{n}(1-x_i)}{1-\theta}$$

$$\frac{\alpha_H}{\theta} = \frac{\alpha_T}{1-\theta} \Rightarrow \alpha_H(1-\theta) = \alpha_T\theta$$

$$\alpha_H = (\alpha_H + \alpha_T)\theta \Rightarrow \theta = \frac{\alpha_H}{\alpha_H+\alpha_T}$$