(6) Demonstrate that the maximum likelihood method for a Bernoulli variable with a linear model is equivalent to minimize the cross-entropy beetween predictions and measured variables.

Bernoulli variable: $\mathcal{L}(p) = \prod_{i}^{n} p^{k_i} (1-p)^{1-k_i}$

Log-Likelihood ($\bar{\mathcal{L}}$) $\bar{\mathcal{L}} := -\ln[\mathcal{L}(p)] = -\ln\left[\prod_{i}^{n} p^{k_i} (1-p)^{1-k_i}\right]$

$\Rightarrow$ using the products property: $\bar{\mathcal{L}} = -\sum_{i=1}^{n} \ln\left[p^{k_i} (1-p)^{1-k_i}\right]$

$$\bar{\mathcal{L}} = -\sum_{i=1}^{n} \left\{ \ln p^{k_i} + \ln\left[(1-p)^{1-k_i}\right] \right\} = -\sum_{i} \left[k_i \ln p + (1-k_i) \ln(1-p)\right]$$

thus, $\operatorname{argmax}(\bar{\mathcal{L}}) = \operatorname{argmin}_{k_i} \left\{ \sum_{i} \left[k_i \ln p + (1-k_i) \ln(1-p)\right] \right\}$

which is the cross-entropy between variables