Assume that samples $\mathcal{X} = \{x_1, \ldots, x_N\}$ follow the Bernoulli distribution with parameter θ, i.e., $p(x) = \theta^x (1-\theta)^{1-x}$. Show that the maximum likelihood estimate of θ is $\theta^{MLE} = \frac{1}{N}\sum_{n=1}^N x_n$.

$l(\theta|\mathcal{X}) = \prod_{n=1}^N p(x_n|\theta) = \prod_{n=1}^N \theta^{x_n}(1-\theta)^{1-x_n}$

$$\mathcal{L}(\theta|\mathcal{X}) = \log l(\theta|\mathcal{X}) = \log \prod_{n=1}^N \theta^{x_n}(1-\theta)^{1-x_n} = \sum_{n=1}^N (x_n \log\theta + (1-x_n)\log(1-\theta))$$
$$= \log\theta \sum_{n=1}^N x_n + \log(1-\theta)\sum_{n=1}^N (1-x_n)$$

We maximize the above expression by setting the first order derivative to zero:

$$\frac{d\mathcal{L}(\theta|\mathcal{X})}{d\theta} = 0 \Rightarrow \frac{1}{\theta}\sum_{n=1}^N x_n - \frac{1}{1-\theta}\sum_{n=1}^N (1-x_n) = 0 \Rightarrow (1-\theta)\sum_{n=1}^N x_n = \theta \sum_{n=1}^N (1-x_n) \Rightarrow$$
$$\sum_{n=1}^N x_n - \theta\sum_{n=1}^N x_n = \theta N - \theta \sum_{n=1}^N x_n \Rightarrow \theta^{MLE} = \frac{1}{N}\sum_{n=1}^N x_n$$

We can further validate that this is a maximum by performing the second order derivative test:

$$\frac{d^2\mathcal{L}(\theta|\mathcal{X})}{d\theta^2} = -\frac{1}{\theta^2}\sum_{n=1}^N x_n - \frac{1}{(1-\theta)^2}\sum_{n=1}^N (1-x_n) < 0 \quad \text{(since } x_n \in [0,1])$$