Working with Likelihoods *(find and work with a partner)*

1. *Bernoulli Random Variable.* Say we flip a weighted coin n times, and each time the probability of heads (1) is p, so the probability of tails (0) is $(1-p)$. Let y_i be the outcome of flip i. For example, if $n = 10$, we might observe these values:

$$\boldsymbol{y} = [0, 0, 1, 1, 0, 1, 0, 1, 0, 0]$$

In this case, the *likelihood* of p given this observed data is

$$L(p) = (1-p) \cdot (1-p) \cdot p \cdot p \cdot (1-p) \cdot p \cdot (1-p) \cdot p \cdot (1-p) \cdot (1-p) = p^4(1-p)^6,$$

since we observe four 1's and six 0's. In general, we can write the likelihood as

$$L(p) = \prod_{i=1}^{n} p^{y_i}(1-p)^{1-y_i} = p^{\#1s}(1-p)^{\#0s},$$

so that for each y_i, only one of y_i and $(1-y_i)$ will be non-zero and contribute to the product. Note that $L(p \mid \boldsymbol{y})$ is a more proper way of writing this (i.e. given the data), but we often omit this conditional part.

(a) What is the *log likelihood* $\ell(p)$ for this setup? Simplify as much as possible.

(b) Our goal is to *maximize* the log likelihood. Take the derivative with respect to p and set it equal to 0. Solve for p – this becomes our MLE (maximum likelihood estimator), $\hat{p}$.

(c) For our concrete example above with $n = 10$, what is the MLE $\hat{p}$? Does this match your intuition?