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 $L(p) = p^4(1-p)^6$, since we observe four 1's and six 0's. In general, we can write the likelihood as

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

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.

We will use the notation $\bar{y} = \frac{1}{n}\sum_{i=1}^{n}$ to denote the mean of $\boldsymbol{y}$.

$$\ell(p) = \log L(p) = \sum_{i=1}^{n} \left[y_i \log(p) + (1-y_i)\log(1-p)\right]$$

$$= \log(p)\sum_{i=1}^{n} y_i + \log(1-p)\sum_{i=1}^{n}(1-y_i)$$

$$= n\bar{y}\log(p) + n(1-\bar{y})\log(1-p)$$

(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), $\hat{p}$.

$$\ell'(p) = \frac{n\bar{y}}{p} - \frac{n(1-\bar{y})}{1-p} = 0 \qquad \text{\# set equal to 0}$$

$$\Rightarrow \cancel{p}(1-p)\frac{\bar{y}}{\cancel{p}} = p\cancel{(1-p)}\frac{(1-\bar{y})}{\cancel{1-p}} \qquad \text{\# divide through by } n \text{ and multiply by } p(1-p)$$

$$\Rightarrow \bar{y} - \bar{y}p = p(1-\bar{y})$$

$$\Rightarrow \boxed{\hat{p} = \bar{y}}$$

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

In our example, we take the mean of $\boldsymbol{y}$:

$$\hat{p} = \frac{0 \cdot 6 + 1 \cdot 4}{10} = \frac{2}{5} = 0.4$$

Although the sample size is small here, this should match our intuition – if we see slightly more tails than heads, then the coin is perhaps weighted slightly towards tails.