In a **logistic regression model**, there are one or more explanatory variables denoted x, and one binary response variable y. The logistic regression model gives a linear function for estimating the log-odds that y is a success, based on the value of x.

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x$$

1. Suppose that the predicted value of the log-odds is 2, that is: $\log\left(\frac{p}{1-p}\right) = 2$. Solve for p. More generally, how do you invert the formula to convert any log-odds L into the probability p?

> **Solution:** If $L = \log(\frac{p}{1-p})$, then
>
> $$\exp(L) = \frac{p}{1-p}$$
>
> After cross-multiplying and collecting like-terms, you should get:
>
> $$p = \frac{\exp(L)}{1+\exp(L)}.$$
>
> So when $L = 2$, $p = 88.1\%$.

The least squares method doesn't work with logistic regression. Instead, we think of each y-value as being a random variable that can be either a success (i.e., $Y = 1$) or a failure ($Y = 0$). This is a simple binomial distribution with $n = 1$. The parameter p in the distribution is: $p = \frac{\exp(\beta_0+\beta_1 x)}{1+\exp(\beta_0+\beta_1 x)}$, so our statistical model is:

$$Y \sim \text{Binom}\left(1, \frac{\exp(\beta_0 + \beta_1 x)}{1 + \exp(\beta_0 + \beta_1 x)}\right)$$

In this model, we think of β_0 and β_1 as the parameters, and x is a fixed value that comes from the data. Recall that the PMF for this binomial distribution is $f(y) = p^y(1-p)^{1-y}$.

2. What is the log-likelihood function for β_0 and β_1 if $y = 1$? What is the log-likelihood function if $y = 0$?

> **Solution:** If $y = 1$:
>
> $$\begin{aligned} \log L(\beta_0, \beta_1 \,|\, y = 1) &= \log(p^y(1-p)^{1-y}) \\ &= \log(p) \\ &= \beta_0 + \beta_1 x - \log(1 + \exp(\beta_0 + \beta_1 x)) \end{aligned}$$
>
> If $y = 0$:
>
> $$\begin{aligned} \log L(\beta_0, \beta_1 \,|\, y = 0) &= \log(p^y(1-p)^{1-y}) \\ &= \log(1-p) \\ &= \log\left(\frac{1}{1 + \exp(\beta_0 + \beta_1 x)}\right) \\ &= -\log(1 + \exp(\beta_0 + \beta_1 x)) \end{aligned}$$

3. What is the log-likelihood function if you have the following three data points? You just need to add the three log-likelihood functions.

x	y	
0	0	← Likelihood function for this data: $L(\beta_0, \beta_1) = -\log(1 + \exp(\beta_0))$
1	0	← Likelihood function for this data: $L(\beta_0, \beta_1) = -\log(1 + \exp(\beta_0 + \beta_1))$
4	1	← Likelihood function for this data: $L(\beta_0, \beta_1) = \beta_0 + 4\beta_1 - \log(1 + \exp(\beta_0 + 4\beta_1))$

Solution:

$$\log L(\beta_0, \beta_1) = -\log(1 + \exp(\beta_0)) - \log(1 + \exp(\beta_0 + \beta_1)) + \beta_0 + 4\beta_1 - \log(1 + \exp(\beta_0 + 4\beta_1))$$

4. Find the two partial derivatives of the log-likelihood function.

Solution:

$$\frac{\partial \log L(\beta_0, \beta_1)}{\partial \beta_0} = -\frac{\exp(\beta_0)}{1 + \exp(\beta_0)} - \frac{\exp(\beta_0 + \beta_1)}{1 + \exp(\beta_0 + \beta_1)} + 1 - \frac{\exp(\beta_0 + 4\beta_1)}{1 + \exp(\beta_0 + 4\beta_1)}$$

$$\frac{\partial \log L(\beta_0, \beta_1)}{\partial \beta_1} = -\frac{\exp(\beta_0 + \beta_1)}{1 + \exp(\beta_0 + \beta_1)} + 4 - 4\frac{\exp(\beta_0 + 4\beta_1)}{1 + \exp(\beta_0 + 4\beta_1)}$$

The equations for when the partial derivatives equal zero are typically too hard to solve without a computer. There are lots of numerical algorithms to find the critical point, however. In the homework, I've given you a problem using R to find the maximum likelihood estimate for β_0 and β_1.