Transformation of parameters: the likelihood

- We will look at the invariance property of MLE in this and the next two videos
- If we transform the parameter from θ to $\phi = \Phi(\theta)$, the value of the likelihood remains the same (provided the transformation is one-to-one)
- This means we can obtain the MLE of ϕ by applying the transformation to θ, i.e. $\hat{\phi} = \Phi(\hat{\theta})$

Setup

Suppose $y_1, y_2, \dots, y_n$ are independent observations with log-likelihood $\ell(\theta; \boldsymbol{Y})$, for a scalar parameter θ.
Consider an invertible, twice differentiable function Φ.
Taking $\phi = \Phi(\theta)$ we can take ϕ as the parameter of interest rather than θ.

e.g. $Y_1, Y_2, \dots, Y_n \sim \text{iid Ber}(\theta)$

log-odds $\phi = \log\left(\frac{\theta}{1-\theta}\right) = \Phi(\theta)$

Relationship between likelihoods

Let the log-likelihoods with respect to $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$ be given respectively by

$$\ell_\theta(\boxed{\theta}; \boldsymbol{y}) \text{ and } \ell_\phi(\phi; \boldsymbol{y}).$$

As Φ is invertible,

$\theta = \Phi^{-1}(\phi)$

It can be checked that the two likelihood are related by

$$\ell_\phi(\phi; \boldsymbol{y}) = \ell_\theta(\Phi^{-1}(\phi); \boldsymbol{y})$$

and

$$\ell_\theta(\theta; \boldsymbol{y}) = \ell_\phi\big(\Phi(\theta; \boldsymbol{y})\big).$$

Example 5.12

Suppose $y_1, y_2, \dots, y_n$ are $i.i.d.$ Bernoulli observations with probability θ.

Consider the log-odds, $\Phi(\theta) = \log\left(\frac{\theta}{1-\theta}\right)$.

Calculate the log-likelihoods for both parameterizations.

Check that $l_\phi(\phi; y) = l_\theta(\Phi^{-1}(\phi); y)$.

θ form

$$p(y_i;\theta) = \theta^{y_i}(1-\theta)^{1-y_i}, \quad y_i = 0, 1$$

$$L(\theta;y) = \prod_{i=1}^{n} p(y_i;\theta) = \prod_{i=1}^{n} \theta^{y_i}(1-\theta)^{1-y_i} = \theta^{\sum_{i=1}^{n} y_i}(1-\theta)^{n-\sum_{i=1}^{n} y_i}$$

$$l(\theta;y) = \left(\sum_{i=1}^{n} y_i\right)\log\theta + \left(n - \sum_{i=1}^{n} y_i\right)\log(1-\theta)$$

ϕ form

$$\phi = \log\left(\frac{\theta}{1-\theta}\right)$$

$$e^\phi = \frac{\theta}{1-\theta}$$

$$(1-\theta)e^\phi = \theta$$

$$e^\phi - \theta e^\phi = \theta$$

$$e^\phi = \theta + \theta e^\phi$$

$$= \theta(1+e^\phi)$$

$$\Rightarrow \theta = \frac{e^\phi}{1+e^\phi} = \Phi^{-1}(\phi)$$

Example 5.12 Solution

$$p(y;\phi) = \left(\frac{e^{\phi}}{1+e^{\phi}}\right)^{y_i}\left(1-\frac{e^{\phi}}{1+e^{\phi}}\right)^{1-y_i} = \left(\frac{e^{\phi}}{1+e^{\phi}}\right)^{y_i}\left(\frac{1}{1+e^{\phi}}\right)^{1-y_i}$$

$$L(\phi;y) = \left(\frac{e^{\phi}}{1+e^{\phi}}\right)^{\sum_{i=1}^{n} y_i}\left(\frac{1}{1+e^{\phi}}\right)^{n-\sum_{i=1}^{n} y_i}$$

$$\ell(\phi;y) = \left(\sum_{i=1}^{n} y_i\right)\log\left(\overset{\Phi^{-1}(\phi)}{\boxed{\frac{e^{\phi}}{1+e^{\phi}}}}\right) + \left(n-\sum_{i=1}^{n} y_i\right)\log\left(\overset{1-\Phi^{-1}(\phi)}{\boxed{\frac{1}{1+e^{\phi}}}}\right)$$

$$= \left(\sum_{i=1}^{n} y_i\right)\log\left(\Phi^{-1}(\phi)\right) + \left(n-\sum_{i=1}^{n} y_i\right)\log\left(1-\Phi^{-1}(\phi)\right)$$

$$= \ell\left(\Phi^{-1}(\phi);y\right)$$

Theorem 15

Suppose $\ell_\theta(\theta; \boldsymbol{y})$ and $\ell_\phi(\phi; \boldsymbol{y})$ are equivalent parametrizations of the same problem. Then

$$\boxed{\hat{\phi} = \Phi(\hat{\theta}).}$$

This property is called the invariance property of MLE.

Proof:

Observe that the possible values of ϕ are $\{\phi : \phi = \Phi(\theta), \theta \in \Theta\}$.

Recall that $\underline{\ell_\phi(\Phi(\theta); y) = \ell_\theta(\theta; y)}$

Maximizing $\ell_\phi(\phi; y)$ wrt ϕ

$\ell_\phi(\Phi(\theta); y)$

$\ell_\theta(\theta; y)$ wrt θ

Hence $\hat{\phi} = \Phi(\hat{\theta})$.

Example 5.13

Suppose $y_1, y_2, \dots, y_n$ are $i.i.d.$ Bernoulli observations with probability θ.

Consider the log-odds, $\Phi(\theta) = \log\left(\frac{\theta}{1-\theta}\right)$.

Find the MLE of θ. What is the MLE of ϕ?

$$\ell(\theta; y) = \left(\sum_{i=1}^n y_i\right)\log\theta + \left(n - \sum_{i=1}^n y_i\right)\log(1-\theta)$$

$$S(\theta; y) = \frac{1}{\theta}\left(\sum_{i=1}^n y_i\right) - \left(\frac{1}{1-\theta}\right)\left(n - \sum_{i=1}^n y_i\right) = 0$$

$$\frac{1}{\theta}\left(\sum_{i=1}^n y_i\right) = \left(\frac{1}{1-\theta}\right)\left(n - \sum_{i=1}^n y_i\right)$$

$$\left(\frac{1-\theta}{\theta}\right)\left(\sum_{i=1}^n y_i\right) = n - \sum_{i=1}^n y_i$$

$$\left(\frac{1}{\theta}-1\right)\left(\sum_{i=1}^n y_i\right) = n - \sum_{i=1}^n y_i$$

$$\frac{1}{\theta}\sum_{i=1}^n y_i \cancel{- \sum_{i=1}^n y_i} = n \cancel{- \sum_{i=1}^n y_i}$$

$$\frac{1}{\theta}\sum_{i=1}^n y_i = n$$

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^n y_i = \bar{y}$$

By Theorem 15,

$$\hat{\phi} = \Phi(\hat{\theta}) = \log\left(\frac{\hat{\theta}}{1-\hat{\theta}}\right) = \log\left(\frac{\bar{y}}{1-\bar{y}}\right)$$

Exercise: Check that the MLE of ϕ derived directly using $\ell_\phi(\phi; \boldsymbol{y})$ is the same as what we have here.

We can check that $\hat{\phi} = \log\left(\frac{\bar{y}}{1-\bar{y}}\right)$ by deriving the MLE of ϕ using the alternative parametrization with ϕ.

$$
\begin{aligned}
&\ell(\phi; \boldsymbol{y}) \\
&= \left(\sum_{i=1}^{n} y_i\right) \log\left(\frac{e^{\phi}}{1+e^{\phi}}\right) + \left(n - \sum_{i=1}^{n} y_i\right) \log\left(\frac{1}{1+e^{\phi}}\right) \\
&= \left(\sum_{i=1}^{n} y_i\right) \log(e^{\phi}) - \left(\sum_{i=1}^{n} y_i\right) \log(1+e^{\phi}) - \left(n - \sum_{i=1}^{n} y_i\right) \log(1+e^{\phi}) \\
&= \left(\sum_{i=1}^{n} y_i\right) \phi - n \log(1+e^{\phi}) \\
&= n\bar{y}\phi - n\log\left(1+e^{\phi}\right)
\end{aligned}
$$

Solving $S(\phi; \boldsymbol{y}) = n\bar{y} - n\left(\frac{e^{\phi}}{1+e^{\phi}}\right) = 0$ for ϕ gives

$$\hat{\phi} = \log\left(\frac{\bar{y}}{1-\bar{y}}\right).$$