Chapter 6 Conditional Likelihoods

6.1 Introduction:

Motivating Example:

2 groups: medicine and placebo.

3 ordinal response categories: no improvement, partial cure, complete cure.

The proportional-odds model,

$$log\left(\frac{r_{ij}}{1-r_{ij}}\right)=\theta_j-x_i\beta, j=1,2, i=1,2,$$

where

$$x_i=\begin{cases}1, medicine\\0, placebo\end{cases}$$

In this example, the parameter β is of interest and θ_1,θ_2 are "nuisance parameters" or "incidental parameter". The number of nuisance parameters might increase as we have more response categories. Thus, the likelihood might depend on a large number of "nuisance parameters" in addition to the parameter of interest.

Objective:

Seek a modified likelihood function that depends on as few of the nuisance parameters as possible while sacrificing as little information as possible.

Let $\theta=(\varphi,\lambda)$, where φ is the parameter vector of interest and λ is a vector of nuisance parameters. The conditional likelihood can be obtained as follows:

1. Find the complete sufficient statistic S_λ.

2. Construct the conditional log-likelihood

$$l_c=log\left(f_{Y|S_\lambda}\right),$$

where $f_{Y|S_\lambda}$ is the conditional distribution of the response

$$Y=\begin{bmatrix}Y_1\\Y_2\\\vdots\\Y_n\end{bmatrix}$$

given S_λ. Two cases might occur. One is that for fixed φ_0, $S_\lambda(\varphi_0)$ depends on φ_0. The other is that $S_\lambda(\varphi_0)=S_\lambda$ is independent of φ_0.

Conditional Likelihood for Exponential Family:

Suppose that the log-likelihood for $\theta=(\varphi,\lambda)$ can be written in the exponential family form

$$l(\theta|y) = \theta^t s - b(\theta).$$

Also, suppose $l(\theta|y)$ has a decomposition of the form

$$l(\theta|y) = \varphi^t s_1 + \lambda^t s_2 - b(\varphi, \lambda)$$

The conditional likelihood of the data Y given s_2 is

$$l_c(\varphi) = l(\varphi|s_2) = \varphi^t s_1 - b^*(\varphi, s_2)$$

which is independent of the nuisance parameter and may be used for inferences regarding φ.

Example 1:

$Y_1 \sim P(\mu_1), Y_2 \sim P(\mu_2)$ are independent. Suppose

$$\varphi = log\left(\frac{\mu_2}{\mu_1}\right) = log(\mu_2) - log(\mu_1)$$

is the parameter of interest and $\lambda_1 = log(\mu_1)$ is the nuisance parameter. Then, the log-likelihood is

$$\begin{aligned}
&l(\varphi, \lambda_1) \\
&\propto log\{exp[-(\mu_1 + \mu_2)] \cdot \mu_1^{y_1} \mu_2^{y_2}\} \\
&= -(\mu_1 + \mu_2) + y_1 \cdot log(\mu_1) + y_2 \cdot log(\mu_2) \\
&= -\mu_1\left(1 + \frac{\mu_2}{\mu_1}\right) + y_1 \cdot log(\mu_1) + y_2 \cdot log(\mu_1) - y_2[log(\mu_1) - log(\mu_2)] \\
&= y_2\varphi + (y_1 + y_2)\lambda_1 - exp(\lambda_1)[1 + exp(\varphi)] \\
&= s_1\varphi + s_2\lambda_1 - b(\varphi, \lambda_1)
\end{aligned}$$

$$\Rightarrow \quad s_1 = y_2, s_2 = y_1 + y_2, b(\varphi, \lambda_1) = exp(\lambda_1)[1 + exp(\varphi)]$$

Then, by the above result for exponential family, the conditional likelihood is

$$l_c(\varphi) = s_1\varphi - b^*(\varphi, s_2).$$

In fact, the conditional distribution of Y_1, Y_2 given $Y_1 + Y_2 = s_2$ is

$$B\left(s_2, \frac{\mu_1}{\mu_1 + \mu_2}\right).$$

Thus,

$$\begin{aligned}
&l_c(\varphi) \\
&\propto y_1 \cdot log\left(\frac{\mu_1}{\mu_1 + \mu_2}\right) + (s_2 - y_1) \cdot log\left(\frac{\mu_2}{\mu_1 + \mu_2}\right) \\
&= y_1 \cdot log\left(\frac{\mu_1}{\mu_1 + \mu_2}\right) + (s_2 - y_1) \cdot log\left(\frac{\mu_1}{\mu_1 + \mu_2}\right) \\
&\quad -(s_2 - y_1)\left[log\left(\frac{\mu_1}{\mu_1 + \mu_2}\right) - log\left(\frac{\mu_2}{\mu_1 + \mu_2}\right)\right] \\
&= y_2\varphi + s_2 \cdot log\left[\frac{1}{1 + exp(\varphi)}\right] \quad \left(by \ \frac{\mu_1}{\mu_1 + \mu_2} = \frac{1}{1 + \left({}^{\mu_2}/_{\mu_1}\right)}\right)
\end{aligned}$$

$$= \boldsymbol{s_1 \varphi - b^*(\varphi, s_2)},$$

where

$$\boldsymbol{b^*(\varphi, s_2) = -s_2 \cdot log\left[\frac{1}{1 + exp(\varphi)}\right]}.$$