Stat 6021: Interpreting Coefficients in Multinomial Logistic Regression

Read this after Section 2 from Guided Notes

1 Binary Logistic Regression

Recall that in the binary logistic regression model, we modeled the log odds as a linear combination of the predictors, i.e.,

$$\text{logit}(\pi) = \log\left(\frac{\pi}{1-\pi}\right) = \boldsymbol{X\beta}, \tag{1}$$

where π denotes the probability of "success", so $1-\pi$ denotes the probability of "failure". Another way to express (1) is to denote the probabilities of "success" and "failure" as π_1 and π_2 respectively. So the logistic regression model in (1) can be written as

$$\log\left(\frac{\pi_1}{\pi_2}\right) = \boldsymbol{X\beta}. \tag{2}$$

The formulation above in (2) is used for multinomial logistic regression.

2 Multinomial Logistic Regression

Suppose there are $m+1$ classes for the response variable. For observation i, consider $m+1$ indicator variables, $Y_{i,1}, \cdots, Y_{i,m+1}$, where:

$$Y_{i,c} = \begin{cases} 1 & \text{if observation } i \text{ is class } c \text{ for } c = 1, \cdots, m+1 \\ 0 & \text{otherwise.} \end{cases}$$

Let $\pi_{i,c}$ denote the probability that observation i belongs to class c for $1 = 2, \cdots, m+1$. Then $\pi_{i,c} = P(Y_{i,c} = 1)$. Using the notation for binary logistic regression in (2), we have

$$\pi'_{i,1,2} = \log\left(\frac{\pi_{i,1}}{\pi_{i,2}}\right) = \boldsymbol{X'_i\beta_{1,2}}.$$

We use $\pi'_{i,1,2}$ and $\boldsymbol{\beta_{1,2}}$ to emphasize that we are modeling the log of the ratio of the probabilities for classes 1 and 2. Suppose we have a response variable with 3 classes. There are 3 pairs of classes, and hence 3 logits, i.e.,

$$
\begin{aligned}
\pi'_{i,1,2} &= \log\left(\frac{\pi_{i,1}}{\pi_{i,2}}\right) = \boldsymbol{X_i'\beta_{1,2}}, \\
\pi'_{i,1,3} &= \log\left(\frac{\pi_{i,1}}{\pi_{i,3}}\right) = \boldsymbol{X_i'\beta_{1,3}}, \\
\pi'_{i,2,3} &= \log\left(\frac{\pi_{i,2}}{\pi_{i,3}}\right) = \boldsymbol{X_i'\beta_{2,3}}.
\end{aligned}
$$

It turns out we only need 2, or 1 less than the number of classes, logits. One class will be chosen as a reference class, and all the other classes will be compared to the reference class, similar to dummy coding. So, supposing class $m+1$ is the baseline class, we need to consider only m comparisons to the reference class. The logit for the cth comparison is

$$\pi'_{i,c,m+1} = \log\left(\frac{\pi_{i,c}}{\pi_{i,m+1}}\right) = \boldsymbol{X_i'\beta_{c,m+1}}$$

for $c = 1, \cdots, m$. Note that $\frac{\pi_{i,c}}{\pi_{i,m+1}}$ is called the **relative risk** of belonging to class c versus belonging to the reference class. Since we consider class $m+1$ to be the baseline, all comparisons are made with class $m+1$, so we let $\pi'_{i,c} = \pi'_{i,c,m+1}$ and $\boldsymbol{\beta_c} = \boldsymbol{\beta_{c,m+1}}$, so the logits are

$$\pi'_{i,c} = \log\left(\frac{\pi_{i,c}}{\pi_{i,m+1}}\right) = \boldsymbol{X_i'\beta_c} \tag{3}$$

for $c = 1, \cdots, m$. We can view (3) as modeling the log relative risk as a linear combination of the predictors.

We only need m logits because the logits for any other comparisons can be obtained from them. For example, suppose $m+1=4$, and we wish to compare classes 1 and 2. We have:

$$
\begin{aligned}
\log\left(\frac{\pi_{i,1}}{\pi_{i,2}}\right) &= \log\left(\frac{\pi_{i,1}}{\pi_{i,4}} \times \frac{\pi_{i,4}}{\pi_{i,2}}\right) \\
&= \log\left(\frac{\pi_{i,1}}{\pi_{i,4}}\right) - \log\left(\frac{\pi_{i,2}}{\pi_{i,4}}\right) \\
&= \boldsymbol{X_i'\beta_1} - \boldsymbol{X_i'\beta_2}.
\end{aligned}
$$

In general, to compare classes k and l, we have

$$\log\left(\frac{\pi_{i,k}}{\pi_{i,l}}\right) = \boldsymbol{X_i'}(\boldsymbol{\beta_k} - \boldsymbol{\beta_l}).$$

The probabilities for each class can be found by

$$\pi_{i,c} = \frac{\exp\left(\boldsymbol{X_i'\beta_c}\right)}{1 + \sum_{k=1}^{m} \exp\left(\boldsymbol{X_i'\beta_k}\right)}$$

for $c = 1, \cdots, m$. The coefficients can be interpreted in one of the following ways:

- The $(k+1)$th element of $\boldsymbol{\beta_c}$ can be interpreted as the **difference in log relative risk** of belonging to class c versus the reference class with a one-unit increase in the kth predictor, given the other predictors are held constant; OR

- The **relative risk ratio** of belonging to class c versus the reference class with a one-unit increase in the kth predictor, given the other predictors are held constant, is the $(k+1)$th element of $\exp(\boldsymbol{\beta_c})$; OR

- The relative risk of belonging to class c versus belonging to the reference class is **multiplied** by the $(k+1)$th element of $\exp(\boldsymbol{\beta_c})$ with a one-unit increase in the kth predictor, given the other predictors are held constant; OR