Applied Statistics With R

Notes on Logistic Regression for R Programming

John Fox

WU Wien
May/June 2006

© 2006 by John Fox
1. Maximum-Likelihood Estimation of the Logistic-Regression Model by Newton-Raphson

The Newton-Raphson method is a common iterative approach to estimating a logistic-regression model:

1. Choose initial estimates of the regression coefficients, such as $b_0 = 0$.

2. At each iteration $t$, update the coefficients:
   \[ b_t = b_{t-1} + (X'V_{t-1}X)^{-1}X'(y - p_{t-1}) \]

   where
   - $X$ is the model matrix, with $x_i'$ as its $i$th row;
   - $y$ is the response vector (containing 0’s and 1’s);

   \[ p_{t-1} = \frac{1}{1 + \exp(-x_i'b_{t-1})} \]

   $V_{t-1}$ is a diagonal matrix, with diagonal entries $p_{i,t-1}(1 - p_{i,t-1})$.

3. Step 2 is repeated until $b_t$ is close enough to $b_{t-1}$. The estimated asymptotic covariance matrix of the coefficients is given by $(X'VX)^{-1}$.


Another approach is to let a general-purpose optimizer do the work of maximizing the log-likelihood,

\[ \log L = \sum y_i \log p_i + (1 - y_i) \log (1 - p_i) \]

- Optimizers work by evaluating the gradient (vector of partial derivatives) of the ‘objective function’ (the log-likelihood) at the current estimates of the parameters, iteratively improving the parameter estimates using the information in the gradient; iteration ceases when the gradient is sufficiently close to zero.

- For the logistic-regression model, the gradient of the log-likelihood is
  \[ \frac{\partial \log L}{\partial b} = \sum (y_i - p_i)x_i \]

- The covariance matrix of the coefficients is the inverse of the matrix of second derivatives. The matrix of second derivatives, called the Hessian, is
  \[ \frac{\partial^2 \log L}{\partial b \partial b'} = X'VX \]

The `optim` function in R, however, calculates the Hessian numerically (rather than using an analytic formula).