Stat 526 Term 2 1998-99 March 9, 1999

Parametric Multiple Regression : the Normal model

Suppose that Y_i, $i = 1, \cdots, n$, are independent normal random variables with common but unknown variance σ^2 and with $\mathrm{E}(Y_i) = \eta(\mathbf{x}_i)$. The log likelihood is

$$\ell(\eta, \sigma^2) = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum (Y_i - \eta(\mathbf{x}_i))^2.$$

Consider a parametric model of η: $\eta(\mathbf{x}) = \eta(\mathbf{x}; \beta_1, \cdots, \beta_{d+p})$ with the β_j's unknown. We can test that a specific submodel holds:

$$H_o : \beta_{p+j} = 0, \quad j = 1, \cdots, d$$

via a generalized log-likelihood ratio test. Let $\hat{\beta}_j^o$, $j = 1, \cdots, p$, and $\hat{\sigma}^{o2}$ be the MLE's under the null model and $\hat{\beta}_j^F$, $j = 1, \cdots, p + d$, and $\hat{\sigma}^{F2}$ be the MLE's under the full model. Note that

$$\hat{\sigma}_o^2 = \frac{1}{n} \sum (Y_i - \eta(\mathbf{x}_i, \hat{\beta}_1^o, \cdots, \hat{\beta}_p^o))^2$$

and

$$\hat{\sigma}_F^2 = \frac{1}{n} \sum (Y_i - \eta(\mathbf{x}_i, \hat{\beta}_1^F, \cdots, \hat{\beta}_{p+d}^F))^2.$$

Two times the generalized log-likelihood ratio statistic is

$$\mathcal{L} = 2\ell(\eta(\cdot, \hat{\beta}_1^F, \cdots, \hat{\beta}_{p+d}^F), \hat{\sigma}_F^2) - 2\ell(\eta(\cdot, \hat{\beta}_1^o, \cdots, \hat{\beta}_p^o), \hat{\sigma}_o^2) = n \ln \left(\frac{\hat{\sigma}_o^2}{\hat{\sigma}_F^2} \right).$$

We reject H_0 if $\mathcal{L}$ is large. To determine the p-value, we can use the fact that, under H_0, $\mathcal{L}$ is asymptotically chi-squared with d degrees of freedom. However, we can also use an exact, nonasymptotic test based on

$$F = \frac{(RSS^o - RSS^F)/d}{RSS^F/(n - p - d)}$$

where $RSS^o = n\hat{\sigma}_o^2$, the residual sum of squares under H_0, and $RSS^F = n\hat{\sigma}_F^2$, the residual sum of squares under the full model. F is large if and only if $\mathcal{L}$ is large. Under H_o, F has an F distribution with degrees of freedom $d, n-p-d$.

Comment: Generalized Likelihood Ratio Test theory does not apply to comparing *non-nested* models, for example to test $H_0 : \beta_{p+j} = 0, j = 1, \cdots, d$, versus $H_A : \beta_i = 0, i = 1, \cdots, p$. However, comparing the RSS's for H_0 and H_A may provide some insight into which is the appropriate model.