4.5. Inference for GLMs

Monday, January 31, 2005
11:25 AM

the likelihood equation

$$\sum_{i=1}^{N} \frac{(y_i - \mu_i) x_{ij}}{\mathrm{Var}(Y_i)} \frac{\partial \mu_i}{\partial \eta_i} = 0$$

y
$\wr\ E(Y)$
μ
$\wr\ g(\mu)$
η
$\wr\ X\beta$
β

($\underset{\sim}{\beta}$ is there implicitly via $\mu_i = g^{-1}(\sum_j \beta_j x_{ij})$)

For most GLMs, it's nonlinear in $\underset{\sim}{\beta}$

Significance testing & interval estimation via

Score
Wald
L-R

) methods apply to any GLM.

4.5.1 Deviance & Goodness of Fit

Let $\tilde{\theta}$ be the estimate of θ for the saturated model (corresponds to $\tilde{\mu}_i = y_i \ \forall i$)

$\hat{\theta}, \hat{\mu}_i$ are ML estimates for a particular unsaturated model.

$$\text{lack of fit} = -2[L(\hat{\underset{\sim}{\mu}}; \underset{\sim}{y}) - L(\underset{\sim}{y}; \underset{\sim}{y})]$$

$$= 2\sum [y_i \tilde{\theta}_i - b(\tilde{\theta}_i)]/a(\phi) - 2\sum_i [y_i \hat{\theta}_i - b(\hat{\theta}_i)]/a(\phi)$$

If $a(\phi) = \frac{\phi}{w_i}$

$$= 2 \sum_i w_i [y_i(\tilde{\theta}_i - \hat{\theta}_i) - b(\tilde{\theta}_i) + b(\hat{\theta}_i)]/\phi$$

$$= \frac{D(\underset{\sim}{y} ; \hat{\underset{\sim}{\mu}})}{\phi} \quad : \text{scaled deviance}$$

deviance

4.5.2 Deviance for Poisson model

$\hat{\theta}_i = \log \hat{\mu}_i \qquad b(\hat{\theta}_i) = \exp(\hat{\theta}_i) = \hat{\mu}_i$

$\tilde{\theta}_i = \log y_i \qquad b(\tilde{\theta}_i) = y_i \qquad a(\phi) = 1$

$$\Rightarrow \quad D(\underset{\sim}{y} ; \hat{\underset{\sim}{\mu}}) = 2 \sum_i y_i \log\left(\frac{y_i}{\hat{\mu}_i}\right)$$

same as G^2 for testing independence

4.5.3 Binomial models

binomial GLMs w/ sample proportions $\{y_i\}$ based on $\{n_i\}$

note
× $y_i \sim \mathrm{bin}(n_i, \pi_i)$
○ $y_i \times n_i \sim \mathrm{bin}(n_i, \pi_i)$
why?

$$\hat{\theta}_i = \log \frac{\hat{\pi}_i}{1 - \hat{\pi}_i}, \quad \tilde{\theta}_i = \log \frac{y_i}{1 - y_i}$$

$$b(\hat{\theta}_i) = \log[1 + \exp(\hat{\theta}_i)] = -\log(1 - \hat{\pi}_i)$$

$\hat{\theta}_i = \log \dfrac{\hat{\pi}_i}{1-\hat{\pi}_i}, \quad \tilde{\theta}_i = \log \dfrac{y_i}{1-y_i}$ why?

$b(\hat{\theta}_i) = \log[1+\exp(\hat{\theta}_i)] = -\log(1-\hat{\pi}_i)$

$b(\tilde{\theta}_i) = -\log(1-y_i)$

$a(\phi) = 1/n_i$ (Thus $\phi = 1$ & $w_i = n_i$)

$$D(\underset{\sim}{y}, \underset{\sim}{\mu}) = 2 \sum \text{observed} \log\left(\frac{\text{observed}}{\text{fitted}}\right)$$

* We can construct data files either as

grouped data	ungrouped data
counts of success at each X setting	individual 0-1 obs. for each subject

* χ^2 approximation holds only for grouped data. (why?)

4.5.4 L-R model comparisons using the deviance

Consider two models $\begin{cases} M_0 & \text{w/ } \hat{\mu}_0 \\ M_1 & \text{w/ } \hat{\mu}_1 \end{cases}$

that are nested, i.e. $M_0 \subset M_1$

simpler

fewer parameters

$\Rightarrow L(\hat{\underset{\sim}{\mu}}_0 ; \underset{\sim}{y}) \le L(\hat{\underset{\sim}{\mu}}_1 ; \underset{\sim}{y})$ why?

smaller maximum

larger deviance

$\Rightarrow D(\underset{\sim}{y} ; \hat{\underset{\sim}{\mu}}_1) \le D(\underset{\sim}{y} ; \hat{\underset{\sim}{\mu}}_0)$

Assuming M_1 holds, testing $H_0: M_0$ uses

$$(*) = D(\underset{\sim}{y} ; \hat{\underset{\sim}{\mu}}_0) - D(\underset{\sim}{y} ; \hat{\underset{\sim}{\mu}}_1)$$

$$= -2\left[L(\hat{\underset{\sim}{\mu}}_0 ; \underset{\sim}{y}) - L(\hat{\underset{\sim}{\mu}}_1 ; \underset{\sim}{y})\right]$$

which is a LR test statistic.

Under regularity conditions,

$$(*) \sim \chi^2(\#\text{ param of } M_1 - \#\text{ param of } M_0)$$

$(*)$ is large if M_0 fits poorly compared to M_1

4.5.5 Residuals for GLMs

If a GLM fits poorly

$\Rightarrow$ examine where the fit is poor by looking at residuals

Note $D(\underset{\sim}{y}; \hat{\underset{\sim}{\mu}}) = \Sigma d_i$ where

$$d_i = 2w_i \left[y_i (\tilde{\theta}_i - \hat{\theta}_i) - b(\tilde{\theta}_i) + b(\hat{\theta}_i) \right]$$

the deviance residual

$\Sigma d_{ij} = G^2$

$$\sqrt{d_i}\ \mathrm{sgn}(y_i - \hat{\mu}_i)$$

The Pearson residual

$$e_i = \frac{y_i - \hat{\mu}_i}{\left[\widehat{\mathrm{var}}(Y_i)\right]^{1/2}}$$

e.g. $\mathrm{var}(Y_i) = \mu_i$ for Poisson

$\Downarrow$

$\Sigma e_{ij}^2 = X^2$ in Pearson χ^2 test of independence

Standardized Poisson residuals

$$r_i = \frac{e_i}{\sqrt{1 - \hat{h}_i}}$$

leverage of obs i.