4.5. Inference for GLMs

The likelihood equation
\[ \sum_{i=1}^{n} \frac{y_i - \mu_i}{\text{var}(y_i)} \mu_i = 0 \]
\( Y \equiv \{ y_i \} \)
\( \mu \equiv \{ \mu_i \} \)
\( \eta \equiv \{ \eta_i \} \)
\( \eta = X \beta \)
\( \beta \)

For most GLMs, it's nonlinear in \( \beta \). Significance testing & interval estimation via score and Wald methods apply to any GLM.

4.5.1 Deviance & Goodness of Fit

Let \( \hat{\Theta} \) be the estimate of \( \Theta \) for the saturated model (corresponds to \( \hat{\mu}; = y_i; x_i \).
\( \hat{\mu}; \) are ML estimates for a particular model.

Unsaturated

\[ \text{deviance} = -2 \left[ L(\hat{\mu};; y_i) - L(\hat{\mu};; y_i) \right] \]
\[ = 2 \sum \frac{y_i \hat{\Theta}_i - b(\hat{\Theta}_i) / a(\hat{\Theta}_i)}{a(\hat{\Theta}_i)} = 2 \sum \frac{y_i (\hat{\Theta}_i - \hat{\Theta}_i) + b(\hat{\Theta}_i)}{a(\hat{\Theta}_i) / \phi} \]
\[ = \frac{D(y_i; \hat{\Theta}_i)}{\phi} \]

Scaled deviance

4.5.2 Deviance for Poisson model
\( \hat{\Theta}_i = \exp(\hat{\lambda}_i); b(\hat{\Theta}_i) = \exp(\hat{\Theta}_i) = \hat{\lambda}_i; \)
\( \hat{\Theta}_i = \exp(y_i); b(\hat{\Theta}_i) = y_i; a(\phi) = 1 \)
\[ = \frac{D(y_i; \hat{\Theta}_i)}{\phi} = 2 \sum y_i \log \left( \frac{y_i}{\hat{\lambda}_i} \right) \sim \text{same as } \chi^2 \text{ for testing independence} \]

4.5.3 Binomial models

Binomial GLMs w/ sample proportions \( \{ y_i \} \)

Based on \( \{ y_i \} \)
\( x_i; y_i \sim \text{bin}(n_i; x_i) \)
\( y_i; x_i; n_i \sim \text{bin}(n_i; y_i) \)
\( \hat{\Theta}_i = \log \frac{\hat{x}_i}{\hat{1} - \hat{x}_i}; \hat{\Theta}_i = \log \frac{\hat{y}_i}{\hat{1} - \hat{y}_i} \)
\( b(\hat{\Theta}_i) = \log [1 + \exp(\hat{\Theta}_i)] = -\log (1 - \hat{\Theta}_i) \)
\[ v_i = \beta + \epsilon_i, \quad \epsilon_i \sim N(0, \sigma^2) \]

* We can construct data files either as grouped data, counts of success at each \( X \) setting, or as individual 0-1 obs. for each subject.

\[ \chi^2 \text{ approximation holds only for grouped data.} \quad \text{(why?)} \]

### 4.5.4 LR model comparisons using the deviance

Consider two models \( M_0, w/ \hat{\beta}_0 \) and \( M_1, w/ \hat{\beta}_1 \) that are nested, i.e., \( M_0 \subset M_1 \).

\[ \Rightarrow L(\hat{\beta}_0; y) \leq L(\hat{\beta}_1; y) \quad \text{why?} \]

\[ \Rightarrow D(y; \hat{\beta}_0) \leq D(y; \hat{\beta}_1) \]

Assuming \( M_1 \) holds, testing \( H_0: M_0 \) vs. \( M_1 \)

\[ * = D(y; \hat{\beta}_0) - D(y; \hat{\beta}_1) = -2 \left[ L(\hat{\beta}_0; y) - L(\hat{\beta}_1; y) \right] \]

which is a LR test statistic. Under regularity conditions,

\[ * \sim \chi^2 \left( \text{# param of } M_1 - \text{# param of } M_0 \right) \]

\( * \) 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 \( D(\mathbf{y}; \mathbf{\hat{\beta}}) = \sum d_i \) where
\[
d_i = 2 w_i \left[ y_i (\hat{\theta}_i - \bar{\theta}_i) - b(\hat{\theta}_i) + b(\bar{\theta}_i) \right]
\]

The deviance residual
\[ \sqrt{d_i} = \text{sgn}(y_i \cdot \hat{\theta}_i) \]

The Pearson residual
\[ e_i = \frac{y_i - \hat{\theta}_i}{\left[ \text{var}(y_i) \right]^{1/2}} \]
\[ \sum e_i^2 = \chi^2 \text{ in Pearson } \chi^2 \text{ test of independence} \]

Standardized Poisson residuals
\[ r_i = \frac{e_i}{\sqrt{1 - h_i}} \quad \text{leverage of obs } i. \]