McCullagh and Nelder

<u>Algorithm (Faraway p. 155)</u>

Iteration number k. Start with k=0.

Initial estimates $\hat{\beta}^{(0)} \Rightarrow \hat{\eta}^{(0)} \Rightarrow \hat{\mu}^{(0)}$

vector of length n

Predicted values

1. Form "adjusted variable"

$$z^{(k)} = \hat{\eta}^{(k)} + \left(y - \hat{\mu}^{(k)}\right) \left.\frac{\partial \eta}{\partial \mu}\right|_{\hat{\eta}^{(k)}}$$

2. New weights

$$\frac{1}{w^{(k)}} = \left.\left(\frac{\partial \eta}{\partial \mu}\right)^2\right|_{\hat{\eta}^{(k)}} \cdot V\left(\hat{\mu}^{(k)}\right)$$

3. Regress $z^{(k)}$ on X with weights $w^{(k)}$

~~4.~~ $\Rightarrow$ New $\hat{\beta}^{(k+1)}$

4. Set $k = k+1$, return to Step 1

Bliss example (p-155)

$$\boxed{\log(1-\mu) = -\log(1+e^{\theta})}$$

$$\boxed{\theta = \log \frac{\mu}{1-\mu}}$$

Binomial Dist^n for Proportions

Let y_i be <u>proportion</u> of successes in

i'th sample of size ~~n~~ n_i

$$f(y_i \mid \mu_i) = \binom{n_i}{n_i y_i} \mu_i^{n_i y_i} (1-\mu_i)^{n_i - n_i y_i}$$

$$= \exp\left[n_i y_i \underbrace{\log\left(\frac{\mu_i}{1-\mu_i}\right)}_{\theta_i} + n_i \log(1-\mu_i) + \log\binom{n_i}{n_i y_i} \right]$$

$$= \exp\left[n_i y_i \theta_i - n_i \log(1+e^{\theta_i}) + \log\binom{n_i}{n_i y_i} \right]$$

$$\eta_i = \theta_i = \log \frac{\mu_i}{1-\mu_i} \qquad \frac{\partial \eta_i}{\partial \mu_i} = \frac{1}{\mu_i} + \frac{1}{1-\mu_i} = \frac{1}{\mu_i(1-\mu_i)}$$

$$V(\mu_i) = \frac{\mu_i(1-\mu_i)}{n_i} \qquad w_i = n_i \mu_i (1-\mu_i)$$

$$\frac{1}{w_i} = \frac{1}{\{\mu_i(1-\mu_i)\}^2} \cdot \frac{\mu_i(1-\mu_i)}{n_i}$$

$$\hat{\text{var}}(\hat\beta) = (X^T W X)^{-1} \hat\phi$$

where W: diagonal entry formed by weights w_i. (Faraway p. 155)

~~Bion~~ Binomial: $\hat\phi = 1$

<u>Hypothesis Tests</u> for Any GLM:

Null model: models without any covariates
[Normal regression: μ = constant)

<u>Full model</u> a.k.a. <u>saturated model</u>
Model that fits data exactly

$$\hat y_i = y_i$$

<u>Deviance</u> for any model <u>against</u> the saturated model

$$D = 2\{ \ell(y, \phi \mid y) - \ell(\hat\mu, \phi \mid y) \}$$

Recall: $\ell(\theta, \phi \mid y) = \frac{\theta y - b(\theta)}{a(\phi)} + c(y, \phi)$ $\quad a(\phi) = \frac{\phi}{w}$

$$D = \frac{2}{\phi} \sum_i w_i \left\{ y_i(\tilde{\theta}_i - \hat{\theta}_i) - b(\tilde{\theta}_i) + b(\hat{\theta}_i) \right\}$$

$\hat{\theta}_i$: estimate under fitted model

$\tilde{\theta}_i$ — — full —

$\tilde{\theta}_i = g^{-1}(y_i)$ in canonical case

Gaussian $w_i = 1$, $\phi = \sigma^2$, $\theta_i = \mu_i$, $b(\theta) = \frac{\theta^2}{2}$

$$D = \frac{2}{\sigma^2} \sum_i \left\{ y_i(y_i - \hat{\mu}_i) - \frac{y_i^2}{2} + \frac{\hat{\mu}_i^2}{2} \right\}$$

$$= \frac{1}{\sigma^2} \sum_i (y_i - \hat{\mu}_i)^2$$

Note σ^2

$\sum (y_i - \hat{\mu}_i)^2 \rightarrow$ Unscaled deviance

$\frac{1}{\sigma^2} \sum (y_i - \hat{\mu}_i)^2 \rightarrow$ Scaled deviance

Tests of Hypotheses

Test 1: Does model fit the data?

$$D \sim \chi^2_{n-p} \quad \text{(approx.)}$$

or: $$X^2 = \sum \frac{(y_i - \hat{\mu}_i)^2}{V(\hat{\mu}_i)}$$ Pearson Statistic

Also $X^2 \sim \chi^2_{n-p}$ (approx)

n: # of obsns

p: # of parameters

Small sample size: χ^2 not a good approx

Logistic regression: doesn't work at all

(in this form)

Tests #2 :

Testing one model against another (nested)

Ω : large model

ω : small model

$\omega \subseteq \Omega$ $\quad$ q : difference in # of parameters

$$D_\omega - D_\Omega \sim \chi^2_q \quad \text{(approx.)}$$

Difficulty : D_ω and D_Ω depend on ϕ and this may be unknown.

One way to estimate ϕ : $\hat{\phi} = \frac{X^2}{n-p}$ $\rightarrow$ Pearson

Relies on : $\frac{X^2}{\phi} \sim \chi^2_{n-p}$

$$\frac{D_\omega - D_\Omega}{q\hat{\phi}} \sim F_{q,n-p} \quad \text{(approx)}$$