Lecture 10

Recap: Likelihood ratio tests

To compare two nested models:

$$G = 2\log\left(\frac{L_{full}}{L_{reduced}}\right)$$

$$= 2(\log L_{full} - \log L_{reduced})$$

$$G \approx \chi^2_q \qquad q = df_{reduced} - df_{full}$$

$$= \#\text{parameters tested}$$

For binary logistic regression: residual deviance = $-2\log L$

$$G = \text{residual deviance}_{reduced} - \text{residual deviance}_{full}$$

What is deviance?

Deviance and dispersion model form

$\underline{\text{EDM}}$: $f(y;\theta,\phi) = a(y,\phi)\exp\left\{\dfrac{y\theta-\kappa(\theta)}{\phi}\right\}$

rewrite in terms of μ, ϕ

Let $t(y,\mu) = y\theta - \kappa(\theta)$

($\theta = g(\mu)$, so $\exists$ some t for which this is true)

Motivation: $N(\mu,\sigma^2)$ $(y-\mu)^2$

$\underline{\text{Claim}}$: $t(y,\mu)$ has a maximum at $\mu = y$

$\underline{\text{Pf}}$: $\dfrac{\partial}{\partial\theta} t(y,\mu) = y - \dfrac{\partial}{\partial\theta}\kappa(\theta) = y-\mu = 0$ when $y=\mu$

$\dfrac{\partial^2}{\partial\theta^2} t(y,\mu) = \dfrac{-\partial^2}{\partial\theta}\kappa(\theta) = \dfrac{-\partial\mu}{\partial\theta} = -V(\mu) < 0$

$\Rightarrow$ $t(y,\mu)$ has a maximum at $y=\mu$ //

$t(y, \mu) = y\theta - \kappa(\theta)$ max. at $y = \mu$

Let $d(y, \mu) = 2\{t(y, y) - t(y, \mu)\}$ <u>unit (unscaled)</u> <u>deviance</u>

$d(y, \mu) = 0$ when $\mu = y$

$d(y, \mu) > 0$ when $\mu \neq y$

$\Rightarrow$ $d(y, \mu)$ measures distance between y and μ

$t(y, y) = y \cdot \theta(y) - \kappa(\theta(y))$

Poisson distribution:

$$f(y;\theta,\phi) = \underbrace{a(y,\phi)}_{\frac{1}{y!}} \underbrace{\exp\left\{\frac{y\theta - \kappa(\theta)}{\phi}\right\}}_{\exp\{y\log\mu - \mu\}}$$

$\phi = 1$

$$t(y,\mu) = y\log\mu - \mu$$

$$t(y,y) = y\log y - y$$

$$d(y,\mu) = 2(y\log y - y - (y\log\mu - \mu))$$

$$= 2\left(y\log\left(\frac{y}{\mu}\right) - (y-\mu)\right)$$

Normal distribution:

$$f(y; \mu, \sigma^2) = \underbrace{\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{y^2}{2\sigma^2}\right\}}_{a(y,\phi)} \exp\left\{\frac{y\mu - \mu^2/2}{\sigma^2}\right\}$$

$$\phi = \sigma^2$$

$$\theta = \mu \qquad \kappa(\theta) = \frac{\mu^2}{2}$$

$$t(y,\mu) = y\theta - \kappa(\theta) = y\mu - \frac{\mu^2}{2}$$

$$t(y,y) = y^2 - \frac{y^2}{2} = \frac{y^2}{2}$$

$$2\left(t(y,y) - t(y,\mu)\right) = 2\left(\frac{y^2}{2} - y\mu + \frac{\mu^2}{2}\right)$$

$$= (y-\mu)^2$$

$$f(y;\theta,\phi) = a(y,\phi)\exp\left\{\frac{y\theta-\kappa(\theta)}{\phi}\right\}$$

$$= a(y,\phi)\exp\left\{\frac{t(y,\mu)}{\phi}\right\}$$

$$= a(y,\phi)\exp\left\{\frac{t(y,\mu)-t(y,y)+t(y,y)}{\phi}\right\}$$

$$= \underbrace{a(y,\phi)\exp\left\{\frac{t(y,y)}{\phi}\right\}}\exp\left\{\frac{t(y,\mu)-t(y,y)}{\phi}\right\}$$

$$f(y;\mu,\phi) = b(y,\phi)\exp\left\{-\frac{d(y,\mu)}{2\phi}\right\}$$

<u>dispersion</u> <u>model</u> <u>form</u>

Residual deviance (unscaled): $D(y,\hat{\mu}) = \sum_{i=1}^{n} d(Y_i,\hat{\mu}_i)$

Scaled residual deviance: $D^*(y,\hat{\mu}) = \dfrac{D(y,\hat{\mu})}{\phi}$

Ex:

$$Y_i \sim N(\mu_i, \sigma^2)$$

$$\mu_i = \beta^T X_i$$

$$d(Y_i, \mu_i) = (Y_i - \mu_i)^2$$

$$D(y, \hat{\mu}) = \hat{\sum_{i=1}} (Y_i - \hat{\mu}_i)^2 = \hat{\sum_{i=1}} (Y_i - \hat{\beta}^T X_i)^2$$

$$D^*(y, \hat{\mu}) = \frac{1}{\sigma^2} \hat{\sum_{i=1}} (Y_i - \hat{\beta}^T X_i)^2$$

$$Y_i \sim EDM(\mu_i, \phi)$$

$$g(\mu_i) = \beta^T X_i \quad \Rightarrow \quad \hat{\mu}_i = g^{-1}(\hat{\beta}^T X_i)$$

$$f(y; \mu, \phi) = b(y, \phi) \exp\left\{ -\frac{d(y, \mu)}{2\phi} \right\}$$

$$L(\hat{\beta}) = \prod_{i=1}^{n} b(Y_i, \phi) \exp\left\{ -\frac{d(Y_i, \hat{\mu}_i)}{2\phi} \right\}$$

$$2 \log L(\hat{\beta}) = 2 \sum_{i=1}^{n} b(Y_i, \phi) - \sum_{i=1}^{n} \frac{d(Y_i, \hat{\mu}_i)}{\phi}$$

$$= 2 \sum_{i=1}^{n} b(Y_i, \phi) - D^*(y, \hat{\mu})$$

$$2 \log L_{full} - 2 \log L_{reduced}$$

$$= \cancel{2 \sum_{i=1}^{n} b(Y_i, \phi) - 2 \sum_{i=1}^{n} b(Y_i, \phi)} + D^*_{reduced}(y, \hat{\mu}) - D^*_{full}(y, \hat{\mu})$$

Likelihood ratio tests when φ is known

Estimating φ

Saddlepoint approximation