STOR 556 3/28/2019

Diagnostics of GLMs
- Residuals
- Influence Diagnostics

Galapagos Example

(a) 3 ways of drawing a residual plot

Linearized Response $z = \eta + (y-\mu)\frac{d\eta}{d\mu}$

$\eta = g(\mu)$

Poisson GLM: $\eta(\mu) = \log\mu$

$\frac{d\eta}{d\mu} = \frac{1}{\mu}$ $\quad z_i = \eta_i + \frac{y_i - \mu_i}{\mu_i}$

(b) Comparing Area and log(Area) as covariates

Partial Residual Plot

Usual definition in linear models

jth covariate (j fixed)

Plot Residuals + $\hat{\beta}_j X_j$ against X_j

X_j: vector of length n

GLM version: $z - \hat{\eta} + \hat{\beta}_j X_j$ versus X_j

Testing Link Functions

Plot linearized response v. linearized predictor

Testing for Outliers

"Half-normal plot"

~~Plot abs~~

Plot ordered ~~absolut~~ absolute residuals

$$v. \ \Phi^{-1}\left(\frac{n+i}{2n+1}\right) \quad i=1, \ldots, n$$

(standard normal. $\Phi^{-1}\left(\frac{i}{n+1}\right)$)

Other examples : Leverage values

Cook's D

Gamma Model
Inverse Gaussian Model

<u>Gamma</u> $\frac{y^{\alpha-1}\beta^{\alpha}e^{-\beta y}}{\Gamma(\alpha)}$ mean $\frac{\alpha}{\beta} \to \mu$

$\nu \equiv \alpha$

$$f(y;\mu,\nu) = \frac{1}{\Gamma(\nu)}\left(\frac{\nu}{\mu}\right)^{\nu} y^{\nu-1} e^{-y\nu/\mu}$$

$$= \exp\left[-\frac{y\nu}{\mu} - \nu\log\mu + \underbrace{\nu\log\nu + (\nu-1)\log y - \log\Gamma(\nu)}_{c(y,\phi)}\right]$$

$\phi = \frac{1}{\nu}$

$\theta = -\frac{1}{\mu}$ $\quad \phi = \frac{1}{\nu}$ $\quad \theta < 0$

$$= \exp\left[\frac{y\theta}{\phi} - \frac{1}{\phi}\log\left(-\frac{1}{\theta}\right) + \cdots\right]$$

$g(\mu) = \frac{1}{\mu}$ $\quad b(\theta) = \log\left(-\frac{1}{\theta}\right) = -\log(-\theta)$

$b'(\theta) = \frac{1}{\theta} = -\mu$

$b''(\theta) = -\frac{1}{\theta^2} = -\mu^2$

variance $\cancel{\phi\mu^2}$ $\phi\, b''(\theta) = \frac{\mu^2}{\nu}$ $\quad\|\quad$ mean $= \frac{\alpha}{\beta}$ $\quad$ var $= \frac{\alpha}{\beta^2}$

$= \left(\frac{\alpha}{\beta}\right)^2 \cdot \frac{1}{\alpha}$

Gamma GLM $\quad b(\theta) = -\frac{1}{\theta}$

$$D(y,\hat{\mu}) = \frac{2}{\phi} \sum W_i \{ y_i(\tilde{\theta}_i - \hat{\theta}_i) - b(\tilde{\theta}_i) + b(\hat{\theta}_i) \}$$

$$= \frac{2}{\phi} \sum_i \left[y_i \left(-\frac{1}{y_i} + \frac{1}{\hat{\mu}_i} \right) + \log\left(\frac{1}{y_i}\right) - \log\left(\frac{1}{\hat{\mu}_i}\right) \right]$$

$$= \frac{2}{\phi} \sum_i \left(- \frac{y_i - \hat{\mu}_i}{\hat{\mu}_i} + \log \frac{y_i}{\hat{\mu}_i} \right),$$

This is Scaled Deviance

Unscaled deviance : omit ϕ

Michaelis–Menten model
relates reaction velocity to concentration
of a substrate

$$E\mu = \mu = \frac{\alpha_0 x}{1+\alpha_1 x}$$

$$\frac{1}{\mu} = \frac{1+\alpha_1 x}{\alpha_0 x} = \frac{\alpha_1}{\alpha_0} + \frac{1}{\alpha_0} \frac{1}{x}$$