Chapter 8 Generalized Linear Models

Exponential family

$$f(y;\theta,\phi) = \exp\left[\frac{y\theta - b(\theta)}{a(\phi)} + c(y,\phi)\right]$$

Density (discrete or continuous) or mass function

y: observation

θ: main parameter of ~~interest~~

ϕ: "dispersion parameter"

$b(\theta)$, $a(\phi)$, $c(y,\phi)$ all functions

Ex 1: Normal

$$f(y;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right]$$

$$= \exp\left[-\frac{y^2}{2\sigma^2} + \frac{y\mu}{\sigma^2} - \frac{\mu^2}{2\sigma^2} - \frac{1}{2}\log(2\pi\sigma^2)\right]$$

$\theta = \mu$, $\phi = \sigma^2$

$$= \exp\left[\frac{y\theta}{\phi} - \frac{(\frac{1}{2}\theta^2)}{\phi} \underbrace{- \frac{y^2}{2\phi} - \frac{1}{2}\log(2\pi\phi)}_{c(y,\phi)} \right]$$

$$b(\theta) = +\frac{1}{2}\theta^2 \qquad a(\phi) = \phi$$

Ex. 2 Poisson $\quad Y \sim \text{Poisson}(\mu)$

$$P(Y=y) = \frac{\mu^y e^{-\mu}}{y!}$$

$$f(y;\mu) = \exp\left[y\log\mu - \mu - \log(y!) \right]$$

$\theta = \log\mu$

$$\exp\left[y\theta - e^\theta - \log(y!) \right]$$

$$b(\theta) = e^\theta \qquad a(\phi) = 1 \qquad c(y,\phi) = -\log(y!)$$

Ex 3 Binomial $\quad y$ as count $\quad \text{Bin}(n, \mu)$

$\uparrow$ probability of success

$$f(y;\mu) = \binom{n}{y} \mu^y (1-\mu)^{n-y}$$

$$= \exp\left[y\log\frac{\mu}{1-\mu} + n\log(1-\mu) + \log\binom{n}{y} \right]$$

Mean is $b'(\theta)$, var $b''(\theta)a(\phi)$

Normal: $b(\theta) = \frac{\theta^2}{2}$ $a(\phi) = \phi = \sigma^2$

$b'(\theta) = \theta = \mu$ $b''(\theta)a(\phi)$ $1 \cdot \sigma^2$

Poisson $b(\theta) = e^{\theta}$ $a(\phi) = 1$

$b'(\theta) = e^{\theta}$ $b''(\theta) = e^{\theta}$

$\mu = e^{\theta}$ Mean = Variance = μ

Binomial $a(\phi) = 1$ $b(\theta) = n\log(1+e^{\theta})$

$$b'(\theta) = n\frac{e^{\theta}}{1+e^{\theta}} = n\left(1 - \frac{1}{1+e^{\theta}}\right)$$

$$b''(\theta) = n\frac{e^{\theta}}{(1+e^{\theta})^2} = \frac{ne^{\theta}}{1+e^{\theta}} \cdot \frac{1}{1+e^{\theta}}$$

$\mu = \frac{e^{\theta}}{1+e^{\theta}}$ Mean $= n\mu$

Var $= n\mu(1-\mu)$

$\theta = \log \frac{\mu}{1-\mu} \qquad e^{\theta} = \frac{\mu}{1-\mu} \qquad \mu = \frac{e^{\theta}}{1+e^{\theta}} \qquad 1-\mu = \frac{1}{1+e^{\theta}}$

$$f(y;\theta) = \exp\left[y\theta - n\log(1+e^{\theta}) + \log \binom{n}{y} \right]$$

$a(\phi) = 1 \qquad b(\theta) = n\log(1+e^{\theta})$

Ex 4 Gamma (p. 175)

$$\frac{\beta^{\alpha} y^{\alpha-1} e^{-\beta y}}{\Gamma(\alpha)} \qquad \frac{\lambda^{\nu} y^{\nu-1} e^{-\lambda y}}{\Gamma(\nu)} \qquad \text{mean } \frac{\nu}{\lambda}$$

Write $\lambda = \frac{\nu}{\mu}$

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

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

$\phi = \frac{1}{\nu} \qquad \theta = -\frac{1}{\mu}$

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

$$b(\theta) = +\log\left(-\frac{1}{\theta}\right)$$

$$= -\log(-\theta)$$

Ex5 Inverse Gaussian $IG(\mu, \lambda)$

$$f(y;\mu,\lambda) = \sqrt{\frac{\lambda}{2\pi y^3}} \exp\left[-\frac{\lambda(y-\mu)^2}{2\mu^2 y}\right]$$

($\lambda \to \infty \Rightarrow$ normal) $\quad y, \mu, \lambda$ all > 0.

Exercise This is an exponential family.

The mean is μ and the variance is $\frac{\mu^3}{\lambda}$

typo in book

Mean and Variance

$$f(y;\theta,\phi) = \exp\left[\frac{y\theta - b(\theta)}{a(\phi)} + c(y,\phi)\right]$$

Log likelihood for a single observation y is

$$\ell(\theta) = \frac{y\theta - b(\theta)}{a(\phi)} + c(y,\phi)$$

$$\ell'(\theta) = \frac{d\ell}{d\theta} = \frac{y - b'(\theta)}{a(\phi)}$$

Fact: $E[l'(\theta); Y] = 0$

$$E[(l'(\theta))^2; Y] = E[-l''(\theta); Y]$$

$$l'(\theta) = \frac{y - b'(\theta)}{a(\phi)} \qquad l''(\theta) = -\frac{b''(\theta)}{a(\phi)}$$

$$0 = E\left[\frac{Y - b'(\theta)}{a(\phi)}\right]$$

$$E\left[\left(\frac{Y - b'(\theta)}{a(\phi)}\right)^2\right] = \frac{b''(\theta)}{a(\phi)}$$

$$E(Y) = b'(\theta) \qquad \text{var } Y = a^2(\phi) \cdot \frac{b''(\theta)}{a(\phi)}$$

$$= b''(\theta)\, a(\phi)$$

we can also write $a(\phi) = \frac{\phi}{w}$

$$\text{var}(Y) = \frac{\phi\, b''(\theta)}{w} \qquad w: \text{weight}$$

Link Functions Mean μ

Some function $\eta = g(\mu)$

Observation i, $\eta_i = \sum_{j=0}^{p} \beta_j x_{ij}$ [usually $x_{i0} = 1$]

Normal: $g(\mu) = \mu$ Poisson: $g(\mu) = \log \mu$

Binomial: $g(\mu) = \log \frac{\mu}{1-\mu}$

g is called link function

Poisson: $\theta = \log \mu$ Binomial $\theta = \log \frac{\mu}{1-\mu}$

If $\theta = \mu$: canonical link function

(mathematical convenience!

Not necessarily best in practice)

Gamma $\theta = -\frac{1}{\mu} < 0$

Usually we would write $\eta = \frac{1}{\mu}$

Fitting a GLM

Sequence of observations y_i, $i = 1, \dots, n$

$E(y_i) = \mu_i$, $\eta_i = g(\mu_i) = \sum_{j=0}^{p} x_{ij}\beta_j$

known

Unknown parameters $\underline{\beta} = (\beta_0, \beta_1, \dots, \beta_p)$

Log likelihood

$$\underline{\ell(\underline{\beta} ; y_i)} = \frac{y_i\theta_i - b(\theta_i)}{a_i(\phi)} + c(y_i ; \phi)$$

$a_i(\phi) = \frac{\phi}{w_i}$

Likelihood equations

$$0 = \sum_{i=1}^{n} \frac{\partial \ell(\beta ; y_i)}{\partial \beta_j} \qquad j = 0, \dots, p$$

$$= \frac{1}{\phi} \sum_i w_i \left[y_i \frac{\partial \theta_i}{\partial \beta_j} - b'(\theta_i) \frac{\partial \theta_i}{\partial \beta_j} \right)$$

Chain rule: $\frac{\partial \mu_i}{\partial \beta_j} = \frac{\partial}{\partial \beta_j}\left(b'(\theta_i)\right) = b''(\theta_i) \frac{\partial \theta_i}{\partial \beta_j}$

$$\frac{\partial l}{\partial \beta_j} = \frac{1}{\phi} \sum_i \frac{y_i - b'(\theta_i)}{b''(\theta_i)/w_i} \cdot \frac{\partial \mu_i}{\partial \beta_j}$$

$$\text{Variance} = b''(\theta)\, a(\phi) = \underbrace{b''(\theta)}_{V}\, \phi$$

$$= \cancel{\sum_i y_i - b'(\theta)}$$

$$= \sum_i \frac{y_i - \mu_i}{V(\mu_i)} \cdot \frac{\partial \mu_i}{\partial \beta_j}$$

Algorithm (Faraway p. 155)

Iteration number k. Start with $k=0$.

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

vector of length n

Predicted values

1. Form "adjusted variable"

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

2. New weights

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

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