Lecture 8

Last time: Poisson regression

Y_i (a count variable $\rightarrow$ values $0, 1, 2, \ldots$)

$Y_i \sim \text{Poisson}(\lambda_i)$ (random component)

$g(\lambda_i) = \beta^T X_i$ (systematic component)

canonical link: $g(\lambda_i) = \log(\lambda_i)$

$$f(y; \lambda) = \frac{e^{-\lambda} \lambda^y}{y!} = \frac{\exp\{y \log \lambda - \lambda\}}{y!}$$

$$= a(y, \phi) \exp\left\{ \frac{y\theta - \kappa(\theta)}{\phi} \right\} \quad \underline{\text{EDM}}$$

$a(y, \phi) = \frac{1}{y!}$, $\theta = \log \lambda$, $\kappa(\theta) = \lambda$, $\phi = 1$

Exponential dispersion models

Bernolli:

$Y \sim \text{Bernolli}(p)$

$$f(y;p) = p^y (1-p)^{1-y}$$

$$= \exp\{\log(p^y(1-p)^{1-y})\}$$

$$= \exp\{\log(p^y) + \log((1-p)^{1-y})\}$$

$$= \exp\{y\log p + (1-y)\log(1-p)\}$$

$$= \exp\{y(\log p - \log(1-p)) + \log(1-p)\}$$

$$= \exp\{y\log\left(\frac{p}{1-p}\right) + \log(1-p)\}$$

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

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

θ is always a function of $E[Y]$

$\phi = 1 \qquad a(y,\phi) = 1$

$$\kappa(\theta) = -\log(1-p)$$

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

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

$$= \log(1+e^\theta)$$

Examples of EDMs

EDM : $f(y; \theta, \phi) = \underbrace{a(y, \phi)}_{\text{normalizing function}} \exp\left\{ \frac{y\theta - \kappa(\theta)}{\phi} \right\}$

- θ is a canonical parameter: function of $E[Y]$

  Say $\mu = E[Y]$, then $\theta = g(\mu)$ (g is a monotonically increasing function)

  ↑ canonical link function

- ϕ is a dispersion parameter is related to $Var(Y)$

  $Var(Y) = \phi \cdot V(\mu)$

Ex: $N(\mu, \sigma^2)$ $Var(Y) = \sigma^2 \cdot 1$ (ϕ, $V(\mu)$)

Bernoulli(p) : $\mu = p$ $Var(Y) = 1 \cdot p(1-p) = 1 \cdot \underbrace{\mu \cdot (1-\mu)}_{V(\mu)}$ (ϕ)

Poisson(λ): $\mu = \lambda$, $Var(Y) = 1 \cdot \mu$ $V(\mu) = \mu$

EDM components

$\kappa(\theta)$ cumulant function

- derivatives of $\kappa(\theta)$ give mean & variance of Y

MGF: $M(t) = \mathbb{E}[e^{tY}]$

$$\mathbb{E}[Y^r] = \frac{d^r}{dt^r} M(t) \Big|_{t=0}$$

Cumulant generating function (CGF):

$$C(t) = \log M(t) = \log \mathbb{E}[e^{tY}]$$

$$\frac{d}{dt} C(t) \Big|_{t=0} = \mathbb{E}[Y] \qquad \text{(HW)}$$

$$\frac{d^2}{dt^2} C(t) \Big|_{t=0} = Var(Y) \qquad \text{(HW)}$$

For an EDM: $M(t) = \mathbb{E}[e^{tY}] = \int_y e^{ty} a(y,\phi)\exp\left\{\frac{y\theta - \kappa(\theta)}{\phi}\right\} dy$

$$= \int_y a(y,\phi)\exp\left\{\frac{y\theta - \kappa(\theta) + ty\phi}{\phi}\right\} dy$$

Idea: try and get $\exp\left\{\frac{y\theta^* - \kappa(\theta^*)}{\phi}\right\}$

$$\theta^* = \theta + t\phi$$

$$\Rightarrow M(t) = \int_y a(y,\phi)\exp\left\{\frac{y\theta^* - \kappa(\theta)}{\phi}\right\} dy$$

$$= \int_y a(y,\phi)\exp\left\{\frac{y\theta^* - \kappa(\theta^*) + \kappa(\theta^*) - \kappa(\theta)}{\phi}\right\} dy$$

$$= \exp\left\{\frac{\kappa(\theta^*) - \kappa(\theta)}{\phi}\right\} \underbrace{\int_y a(y,\phi)\exp\left\{\frac{y\theta^* - \kappa(\theta^*)}{\phi}\right\} dy}_{1}$$

$$= \exp\left\{\frac{\kappa(\theta^*) - \kappa(\theta)}{\phi}\right\} = \exp\left\{\frac{\kappa(\theta + t\phi) - \kappa(\theta)}{\phi}\right\}$$

$$\Rightarrow C(t) = \frac{\kappa(\theta + t\phi) - \kappa(\theta)}{\phi}$$

$$c(t) = \frac{\kappa(\theta + t\phi) - \kappa(\theta)}{\phi}$$

Conclusion: $\frac{d\kappa(\theta)}{d\theta} = \mu$

$\frac{d\mu}{d\theta} = V(\mu)$

$$\frac{d\,c(t)}{dt} = \frac{\kappa'(\theta + t\phi)}{\phi} \cdot \phi = \kappa'(\theta + t\phi)$$

$$\underbrace{\frac{d\,c(t)}{dt}\Big|_{t=0}}_{=\mathbb{E}[Y]} = \kappa'(\theta) = \frac{d\kappa(\theta)}{d\theta}$$

<u>So</u> $\frac{d\kappa(\theta)}{d\theta} = \mathbb{E}[Y] = \mu$

$$\underbrace{\frac{d^2 c(t)}{dt^2}\; c(t)\Big|_{t=0}}_{=\mathrm{Var}(Y)} = \phi \cdot \frac{d^2 \kappa(\theta)}{d\theta^2}$$

<u>So</u> $\mathrm{Var}(Y) = \phi \cdot \frac{d^2}{d\theta^2}\kappa(\theta) = \phi \cdot \frac{d}{d\theta}\left(\frac{d\kappa(\theta)}{d\theta}\right)$

$$= \phi \cdot \frac{d}{d\theta}\mu(\theta) = \phi \cdot V(\mu)$$

Cumulants and the cumulant generating function

Generalized linear models