{show ... }

☐ GLIM:

$$f(y \mid \theta_i, \phi) = \exp\left\{ [y\,\theta_i - b(\theta_i)] / \left(\frac{\phi}{w_i}\right) + c_i(y, \phi) \right\}$$

where ① $g(\mu_i) = \eta_i = \underset{\sim}{x}_i^t \underset{\sim}{\beta} + c_i$ (def. of GLIM)

and ② $b'(\theta_i) = \mu_i$ } for exponential family

③ $\left(\frac{\phi}{w_i}\right) b''(\theta_i) = var(y_i)$

☐ So for independent $y_1, \ldots, y_n$

$$L = \sum_{i=1}^{n} \left\{ \frac{[y_i\theta_i - b(\theta_i)]}{\phi/w_i} + c_i(y_i, \phi) \right\}$$

$\Rightarrow$

(*) $$\frac{\partial L}{\partial \beta_j} = \frac{1}{\phi} \sum_{i=1}^{n} w_i [y_i - b'(\theta_i)] \frac{\partial \theta_i}{\partial \beta_j}$$

$\{f'g + fg'\}$

and $$\frac{\partial^2 L}{\partial \beta_k \partial \beta_j} = \frac{1}{\phi} \sum_{i=1}^{n} w_i \frac{\partial}{\partial \beta_k} \left\{ \underbrace{[y_i - b'(\theta_i)]}_{f} \cdot \underbrace{\frac{\partial \theta_i}{\partial \beta_j}}_{g} \right\}$$

$$= \frac{1}{\phi} \sum_{i=1}^{n} w_i \left\{ - \frac{\partial}{\partial \beta_k} [b'(\theta_i)] \cdot \frac{\partial \theta_i}{\partial \beta_j} + [y_i - b'(\theta_i)] \frac{\partial^2 \theta_i}{\partial \beta_k \partial \beta_j} \right\}$$

(**) $$= \frac{1}{\phi} \sum_{i=1}^{n} w_i \left\{ -b''(\theta_i) \left(\frac{\partial \theta_i}{\partial \beta_k}\right) \left(\frac{\partial \theta_i}{\partial \beta_j}\right) + [y_i - b'(\theta_i)] \frac{\partial^2 \theta_i}{\partial \beta_k \partial \beta_j} \right\}$$

☐ NEED $b'(\theta_i)$, $b''(\theta_i)$, $\frac{\partial \theta_i}{\partial \beta_j}$, $\frac{\partial^2 \theta_i}{\partial \beta_k \partial \beta_j}$ {in terms of μ_i, σ_i^2, x_{ij}}

☐ Using

$$① \quad g(\mu_i) = \underset{\sim}{x}_i^t \underset{\sim}{\beta} + c_i$$

$$② \quad b'(\theta_i) = \mu_i$$

$$③ \quad \left(\frac{\phi}{w_i}\right) b''(\theta_i) = var(y_i)$$

gives

☐ $b'(\theta_i) = \mu_i$

$$b''(\theta_i) = \frac{w_i}{\varphi} \cdot var(y_i)$$

☐ For $\frac{\partial \theta_i}{\partial \beta_j}$, consider $\theta_i \longleftarrow \mu_i \longleftrightarrow \beta_j$.

$$\Rightarrow \frac{\partial \theta_i}{\partial \beta_j} = \overbrace{\frac{d\theta_i}{d\mu_i}}^{(a)} \cdot \overbrace{\frac{\partial \mu_i}{\partial \beta_j}}^{(b)}$$

(a) $b'(\theta_i) = \mu_i \Rightarrow b''(\theta_i)\, d\theta_i = d\mu_i$

$$\Rightarrow \frac{d\theta_i}{d\mu_i} = \frac{1}{b''(\theta_i)}$$

(b) $g(\mu_i) = \underset{\sim}{x}_i^t \underset{\sim}{\beta} + c_i$

$$\Rightarrow g'(\mu_i) \frac{\partial \mu_i}{\partial \beta_j} = x_{ij}$$

$$\Rightarrow \frac{\partial \theta_i}{\partial \beta_j} = \underbrace{\frac{1}{b''(\theta_i)}}_{(a)} \cdot \underbrace{\frac{1}{g'(\mu_i)} x_{ij}}_{(b)}$$

$$\overset{③}{=} \frac{\phi}{w_i \cdot var(y_i)} \cdot \frac{1}{g'(\mu_i)} \cdot x_{ij}$$

☐ For $\dfrac{\partial^2 \theta_i}{\partial \beta_k \partial \beta_j}$: Too complicated

(15.1)

{will see later that sometimes can do without}

☐ Using

$\boxed{b'(\theta_i)} = \mu_i \qquad \boxed{b''(\theta_i)} = \dfrac{w_i}{\varphi} \cdot var(Y_i) \qquad \boxed{\dfrac{\partial \theta_i}{\partial \beta_j}} = \dfrac{\varphi \cdot x_{ij}}{w_i \cdot var(Y_i) \cdot g'(\mu_i)}$

gives (from ⊛)

$$\frac{\partial L}{\partial \beta_j} = \frac{1}{\varphi} \sum_{i=1}^{n} w_i \left[y_i - b'(\theta_i)\right] \frac{\partial \theta_i}{\partial \beta_j}$$

$$= \frac{1}{\varphi} \sum_{i=1}^{n} w_i [y_i - \mu_i] \cdot \frac{\varphi\, x_{ij}}{w_i\, var(Y_i)\, g'(\mu_i)} = \sum_{i=1}^{n} x_{ij} \frac{(y_i - \mu_i)}{g'(\mu_i)\, var(Y_i)}$$

☐ and (from ⊛⊛)

$$\frac{\partial^2 L}{\partial \beta_k \partial \beta_j} = \frac{1}{\varphi} \sum_{i=1}^{n} w_i \left\{ -b''(\theta_i) \left(\frac{\partial \theta_i}{\partial \beta_k}\right)\left(\frac{\partial \theta_i}{\partial \beta_j}\right) + [y_i - b'(\theta_i)] \frac{\partial^2 \theta_i}{\partial \beta_k \partial \beta_j} \right\}$$

$$= \frac{1}{\varphi} \sum_{i=1}^{n} w_i \left\{ -\frac{w_i\, var(Y_i)}{\varphi} \cdot \frac{\varphi\, x_{ij}}{w_i\, var(Y_i)\, g'(\mu_i)} \cdot \frac{\varphi\, x_{ik}}{w_i\, var(Y_i)\, g'(\mu_i)} + [y_i - \mu_i] \frac{\partial^2 \theta_i}{\partial \beta_k \partial \beta_j} \right\}$$

$$= \sum_{i=1}^{n} - \frac{x_{ij}\, x_{ik}}{var(Y_i)[g'(\mu_i)]^2} + \underbrace{\frac{1}{\varphi} \sum_{i=1}^{n} w_i (y_i - \mu_i) \cdot \frac{\partial^2 \theta_i}{\partial \beta_k \partial \beta_j}}_{\text{II}}$$

☐ {See formula page}

☐ Suppose can ignore II in $\frac{\partial^2 L}{\partial \beta_k \partial \beta_j}$. Then

$$\frac{\partial L}{\partial \beta_j} = \sum_{i=1}^{n} x_{ij} \cdot \frac{1}{var(Y_i)} \cdot \frac{(y_i - \mu_i)}{g'(\mu_i)}$$

and

$$\frac{\partial^2 L}{\partial \beta_k \partial \beta_j} = - \sum_{i=1}^{n} x_{ik} \cdot \frac{1}{var(Y_i)[g'(\mu_i)]^2} \cdot x_{ij}$$

skip

☐ In matrix form:

$$\left(\frac{\partial^2 L}{\partial \beta_k \partial \beta_j} \right)_{p \times p} = - X^t A^{-1} X, \quad A = diag\left(var(Y_i)[g'(\mu_i)]^2\right)$$

$$\frac{\partial L}{\partial \beta_j} = \sum_{i=1}^{n} x_{ij} \frac{1}{var(Y_i)[g'(\mu_i)]^2} \cdot g'(\mu_i)(y_i - \mu_i)$$

so

$$(\nabla L)_{p \times 1} = X^t A^{-1} \underset{\sim}{r} \quad \text{for} \quad r_i = \underbrace{g'(\mu_i)(y_i - \mu_i)}_{\text{"working residual"}}$$