Data sets for project

- will get subsets by email
  (NEED TO REQUEST)

- Description:

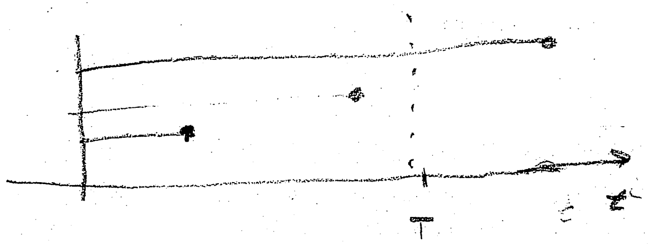
via email

Birthweights — Bernoulli

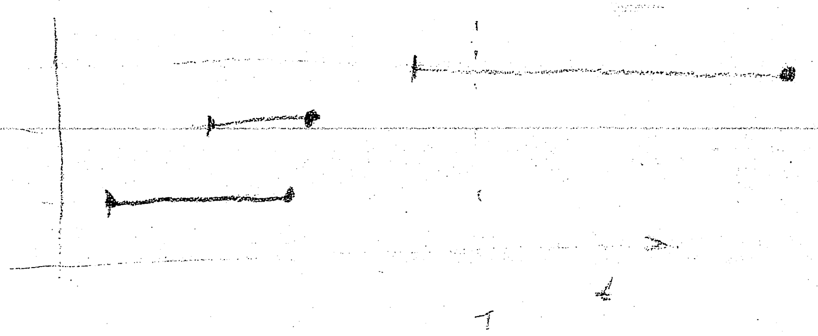
Shock data — Bernoulli

rat sightings — Poisson / negative binomial
— spatial data

No Gamma example — couldn't find large data set w/out censoring, survival analysis



controlled



observational study

□ Def of GLIM

(1) $Y_1, \ldots, Y_n$ indep, $E(Y_i) = \mu_i$

$\underset{\sim}{x}_1, \ldots, \underset{\sim}{x}_n$

(2) Y_i : exponential family

Density (prob) fn

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

where

b, c_i : known functions

w_i : known constants (wts); sometimes $w_i \equiv 1$

θ_i : unknown parameters (natural parameters)

ϕ : $\begin{cases} \text{unknown parameter (scale parameter)} \\ = 1 \end{cases}$

(3) Relation between μ_i and $\underset{\sim}{x}_i$:

For monotonic, increasing <u>link function</u> g,

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

for known constants (off sets) c_i (possibly $c_i \equiv 0$)

□ Q: How connect between μ_i in (3) and θ_i in (2)?

☐ A: Recall that for exponential family,

(A) $E(Y_i) = \mu_i = b'(\theta_i)$

(B) $var(Y_i) = \frac{\phi}{w_i} \cdot b''(\theta_i)$

{ Proof: $\int f(y|\theta)\,dy = 1$ for all θ

$\Rightarrow \frac{\partial}{\partial\theta}\int f(y|\theta)\,dy = \frac{\partial^2}{\partial\theta^2}\int f(y|\theta)\,dy = 0.$

Then interchange order of $\frac{\partial}{\partial\theta}$ and $\int$;

details on next page.

☐ ϕ is called the <u>scale parameter</u>

θ_i ———————— <u>location parameter</u>

$$f(y|\theta,\phi) = e^{[y\theta - b(\theta)]/(\frac{\phi}{w}) + c(y,\phi)}$$

① $$0 = \frac{\partial}{\partial\theta}\int f(y|\theta)\,dy = \int \underbrace{\frac{\partial}{\partial\theta} f(y|\theta)}\,dy$$

$$= \int \frac{\partial}{\partial\theta}\left\{\frac{[y\theta - b(\theta)]}{(\phi/w)} + c(y,\phi)\right\} f(y)\,dy \quad (**)$$

$$= \int \left[\frac{y - b'(\theta)}{(\phi/w)} + 0\right] f(y)\,dy = \underbrace{\frac{w}{\phi}\int [y - b'(\theta)]\, f(y)\,dy}_{(*)}$$

$$\Rightarrow \quad E(Y) = b'(\theta).$$

② $$0 = \frac{\partial^2}{\partial\theta^2}\int f(y|\theta)\,dy \overset{(*)}{=} \frac{\partial}{\partial\theta}\,\frac{w}{\phi}\int \overbrace{[y - b'(\theta)]}^{f}\,\overbrace{f(y|\theta,\phi)}^{g}\,dy$$

$$\Rightarrow \quad 0 = \int \overbrace{\frac{\partial}{\partial\theta}[y - b'(\theta)]}^{f'}\cdot \overbrace{f(y|\theta,\phi)}^{g}\,dy + \int [y - b'(\theta)]\,\underbrace{\frac{\partial}{\partial\theta} f(y|\theta,\phi)}_{(**)}\,dy$$

$$= \int -b''(\theta)\, f(y|\theta)\,dy + \int (y-\mu)\,\frac{(y-\mu)}{\phi/w}\cdot f(y|\theta,\phi)\,dy$$

$$= -b''(\theta) + \frac{w}{\phi}\,\mathrm{var}(Y)$$

$$\Rightarrow \quad \mathrm{var}(Y) = \frac{\phi}{w}\, b''(\theta)$$

Proof that for the exponential family,

$$E(Y) = b'(\theta)$$

$$\mathrm{var}(Y) = \frac{\phi}{w}\, b''(\theta).$$

Example: θ_i, b, w_i, ϕ and c

for $y_i \sim N(\mu_i, \sigma^2)$.

$$f(y) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(y-\mu_i)^2}{2\sigma^2}}$$

exp. family $\downarrow$

$$\left\{ \log f_i(y) = \left[\frac{y\theta_i - b(\theta_i)}{\phi}\right] \cdot w_i + c_i(y, \phi) \right.$$

$\Rightarrow$

$$\log f(y) = \frac{-(y-\mu_i)^2}{2\sigma^2} - \log(\sqrt{2\pi}\,\sigma)$$

$$= \frac{-y^2 + 2y\mu_i - \mu_i^2}{2\sigma^2} - \log(\sigma) - \log(\sqrt{2\pi})$$

$$= \frac{2y\mu_i - \mu_i^2}{2\sigma^2} \underbrace{- \frac{y^2}{2\sigma^2} - \log\sigma - \log(\sqrt{2\pi})}$$

$$= \frac{y \cdot \mu_i - \mu_i^2/2}{\sigma^2} + \downarrow$$

Thus $\theta_i = \mu_i$

$b(\theta_i) = \theta_i^2/2$ $\quad \mu_i^2/2$

$\phi = \sigma^2$, $w_i \equiv 1$, $c(y, \phi) = c(y, \sigma^2) =$

(A): $b'(\theta_i) = \theta_i = E(y_i)$

(B) $b''(\theta_i) = 1 \Rightarrow \frac{\phi}{w_i} b''(\theta_i) = \sigma^2 = Var(y_i)$

□ Q: How compute MLE for GLIM? 12.1

EG: $Y_i \sim \text{Pois}(\lambda_i)$

$$\log(\lambda_i) = x_i \beta \quad \{x_i \text{ scalar}\}$$

$$f(y_1, \ldots, y_n) = \prod_{i=1}^{n} \frac{(\lambda_i)^{y_i} e^{-\lambda_i}}{y_i!}$$

$\Rightarrow$

$$L = \sum_{i=1}^{n} [y_i \log(\lambda_i) - \lambda_i] + \text{constant}$$

$$= \sum_{i=1}^{n} [y_i x_i \beta - e^{x_i \beta}]$$

□ $$\frac{dL}{d\beta} = \left(\sum_{i=1}^{n} y_i x_i\right) - \sum_{i=1}^{n} x_i e^{x_i \beta} = 0$$

$$\Leftrightarrow \quad \sum_{i=1}^{n} x_i \cdot e^{\beta x_i} = \sum_{i=1}^{n} x_i y_i$$

□ Solution? e.g., $2e^{\beta} + e^{2\beta} = 17$?

$$\left[x_i = 1 : \quad \sum_{i=1}^{n} x_i e^{\beta x_i} = e^{\beta} \cdot n = \sum_{i=1}^{n} y_i \right.$$

$$\left. \Rightarrow \hat{\beta} = \log(\bar{y}) \right]$$

□ So need to use NR

□ MLE of $\underset{\sim}{\beta}$ using NR algorithm

To solve $\nabla_{\underset{\sim}{\beta}} L = \begin{bmatrix} \frac{\partial}{\partial \beta_1} L \\ \vdots \\ \frac{\partial}{\partial \beta_p} L \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$ for GLIM,

where L = log-lik for GLIM

□ NR: To solve

$$\underset{\sim}{h}(\underset{\sim}{x}) = \begin{bmatrix} h_1(\underset{\sim}{x}) \\ \vdots \\ h_p(\underset{\sim}{x}) \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix},$$

use

$$\underset{\sim}{x}_{(m+1)} = \underset{\sim}{x}_{(m)} - \left(\frac{\partial h_j}{\partial x_k} \right)^{-1}_{p \times p} \cdot \underset{\sim}{h}(\underset{\sim}{x}_{(m)})$$

So

(*) $$\underset{\sim}{\beta}_{(m+1)} = \underset{\sim}{\beta}_{(m)} - \left(\frac{\partial^2 L}{\partial \beta_j \partial \beta_k} \right)^{-1} \begin{bmatrix} \frac{\partial L}{\partial \beta_1}(\underset{\sim}{\beta}_{(m)}) \\ \vdots \\ \frac{\partial L}{\partial \beta_p}(\underset{\sim}{\beta}_{(m)}) \end{bmatrix}$$

□ So need $\frac{\partial L}{\partial \beta_j}$ and $\frac{\partial^2 L}{\partial \beta_j \partial \beta_k}$

[] 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 $b'(\theta_i) = \mu_i$

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

[] 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 $b'(\theta_i)$, $b''(\theta_i)$,

$$\frac{\partial \theta_i}{\partial \beta_j}, \quad \frac{\partial^2 \theta_i}{\partial \beta_k \partial \beta_j}$$

□ Using

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

② $b'(\theta_i) = \mu_i$

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

gives

□ $b'(\theta_i) = \mu_i$ ✓

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

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

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

(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

{will see later that sometimes can do without}

Using

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

gives

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

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

and

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

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

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

{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

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

☐ In matrix form:

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

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

so

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

In matrix form:

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

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

$\Rightarrow$

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

[Normal dist

Thus NR: $\underset{\sim}{\beta}_{(m+1)} = \underset{\sim}{\beta}_{(m)} - \left(\frac{\partial^2 L}{\partial \beta_j \partial \beta_k}\right)^{-1} \underset{\sim}{\nabla} L(\underset{\sim}{\beta}_{(m)})$

becomes

$$(*) \qquad \underset{\sim}{\beta}_{(m+1)} = \underset{\sim}{\beta}_{(m)} + (X^t A_m^{-1} X)^{-1} X^t A_m^{-1} \underset{\sim}{r}_{(m)}$$

where A_m, $\underset{\sim}{r}_{(m)}$ evaluated for $\underset{\sim}{\beta} = \underset{\sim}{\beta}_{(m)}$

(*) is example of IRWLS

Iteratively Reweighted Least Squares,

$\boxed{\text{SEE FORMULA PAGE}}$