SIMPLE LINEAR REGRESSION OPENBUGS EXAMPLE

$y_i \sim N(\mu_i, \sigma^2)$ WHERE $\tau = 1/\sigma^2$ PRECISION PARAMETER

$\mu_i = \alpha + \beta(x_i - \bar{x})$

BAYES APPROACH

$\alpha \sim N(0, 0.001)$, $\beta \sim N(0, 0.001)$

NORMAL PRIORS WITH VARIANCE $1/0.001 = 1000$

$\tau \sim \text{GAMMA}(0.1, 0.1)$ MEAN OF $\tau = 1$ VAR OF $\tau = 10$

2ND MODEL

$y \sim \text{BINOMIAL}(n, \theta)$ $n = 100$ TRIALS $Y = 10$

$\text{logit}(\theta) = \log\left(\frac{\theta}{1-\theta}\right) = \psi$

$\psi \sim N\left(0, \frac{1}{0.368}\right) \equiv N(0, 2.71)$

$\Rightarrow P\left(-2 < \frac{\psi}{\sqrt{2.71}} < 2\right) = 0.95 \Rightarrow P(-3.29 < \psi < 3.29) = 0.95$

IF $\psi \approx 0$

$\Rightarrow \log\left(\frac{\theta}{1-\theta}\right) \approx 0 \Rightarrow \frac{\theta}{1-\theta} \approx 1 \Rightarrow \theta \approx 0.5$

MODEL EXAMPLE TABLE 2.7

$Y_{jk} \sim N(\mu_j, \sigma^2)$; $j = 1, 2$, $k = 1, 2, \ldots 20$

$E(Y_{jk}) = \mu_j$, $j = 1, 2$

THE NULL MODEL $\mu_1 = \mu_2 \Rightarrow E(Y_{jk}) = \mu$ (ASSUME σ^2 KNOWN)

FIND THE MAXIMUM LIKELIHOOD ESTIMATOR UNDER THE NULL MODEL

$f(y_{jk} \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(y_{jk} - \mu)^2\right)$

$$L(\mu, \sigma^2) = \prod_{j=1}^{2} \prod_{k=1}^{20} \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) \exp\left(-\frac{1}{2\sigma^2}(y_{jk} - \mu)^2\right)$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^{20 \times 2} \exp\left(-\frac{1}{2\sigma^2} \sum_{j=1}^{2} \sum_{k=1}^{20} (y_{jk} - \mu)^2\right)$$

$$\log L(\mu) = -\frac{1}{2\sigma^2}\sum_{j=1}^{2}\sum_{k=1}^{20}(y_{jk}-\mu)^2 + \text{CONSTANT}$$ SO FINDING MAX EQUIV. TO MIN SUM OF SQUARES

$$\frac{d\log L(\mu)}{d\mu} = +\frac{2}{2\sigma^2}\sum_{j=1}^{2}\sum_{k=1}^{20}(y_{jk}-\mu) = 0 \qquad \frac{d}{d\mu}(y_{jk}-\mu)^2 = -(y_{jk}-\mu)$$

$$\Rightarrow \sum_{j=1}^{2}\sum_{k=1}^{20} y_{jk} = (20)(2)\mu \quad \text{OR} \quad \hat{\mu} = \bar{Y}_{..} = \left(\sum_{j=1}^{2}\sum_{k=1}^{20} y_{jk}\right)\Big/(20)(2)$$

UNDER H_1:

$$\log L(\mu_1,\mu_2) = -\frac{1}{2\sigma^2}\left[\sum_{k=1}^{20}(y_{1k}-\mu_1)^2 + \sum_{k=1}^{20}(y_{2k}-\mu_2)^2\right]$$

$$\frac{d\log L(\mu_1,\mu_2)}{d\mu_j} = \frac{1}{\sigma^2}\sum_{k=1}^{20}(y_{jk}-\mu_j) = 0 \Rightarrow \hat{\mu}_j = \bar{Y}_{j.} = \sum_{k=1}^{20} y_{jk}/20$$

REMARKS ON NOTATION

$$g(E(y)) = \beta_0 + \beta_1 X_1 + \dots + \beta_m X_m \quad \text{OR}$$
$$g(E(Y)) = X\beta$$

$g(E(Y))$ IS THE VECTOR OF MEANS $g(E(Y))^t = (g(E(Y_1)), \dots g(E(Y_n)))$

$X_{n\times p}$ MATRIX COLS. ARE COVARIATES AND LEVELS.

POISSON MODEL FOR TWO GROUPS

$E(Y_{jk}) = \theta_j$, $Y_{jk} \sim \text{Poisson}(\theta_j)$ $j=1,2$

$$Y = \begin{bmatrix} Y_{11} \\ Y_{12} \\ \vdots \\ Y_{1,20} \\ Y_{21} \\ \vdots \\ Y_{2,23} \end{bmatrix}; \quad g(\theta_j) = \theta_j, \quad \beta = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}_{2\times 1}; \quad X = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 0 & 1 \\ \vdots & \vdots \\ 0 & 1 \end{bmatrix}_{49\times 2}$$

$$X\beta = \begin{bmatrix} \theta_1 \\ \theta_1 \\ \vdots \\ \theta_1 \\ \theta_2 \\ \vdots \\ \theta_2 \end{bmatrix}_{49\times 1}$$

$Y_{jk} \sim N(\mu_j, \sigma^2)$ $\quad g(\mu_j) = \mu_j$

$$\beta = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}; \quad X = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 0 & 1 \\ \vdots & \vdots \\ 0 & 1 \end{bmatrix}_{40\times 2}$$

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

$E(y_i) = \alpha + \beta(x_i - \bar{x}) = \mu_i$

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}_{n\times 1} \qquad \underbrace{g(\mu_i)}_{\text{LINK FUNCTION}} \overset{=\mu_i}{=} \alpha + \beta(x_i - \bar{x}) \;,\; \underline{\beta} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{2\times 1} ; \; X = \begin{bmatrix} 1 & x_1 - \bar{x} \\ 1 & x_2 - \bar{x} \\ \vdots & \\ 1 & x_n - \bar{x} \end{bmatrix}_{n\times 2}$$

$$X\beta = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix}$$

$Y \sim$ BINOMIAL(n, θ) $\quad E(Y) = n \cdot \theta$

BUT $\log\left(\frac{\theta}{1-\theta}\right) = \varphi = g(\theta)$ FOCUS ON φ (NO REGRESSOR ONLY INTERCEPT)

EXTENSION TO LOGISTIC REGRESSION $\log\left(\frac{\theta}{1-\theta}\right) = \alpha + \beta x$

$$f(y;\theta) = \binom{n}{y} \theta^y (1-\theta)^{n-y} = \binom{n}{y} \left(\frac{\theta}{1-\theta}\right)^y (1-\theta)^n$$

$$= \binom{n}{y} \exp\left(y \ln\left(\frac{\theta}{1-\theta}\right)\right) (1-\theta)^n$$

$$= \exp\left(y \ln\left(\frac{\theta}{1-\theta}\right) + n\log(1-\theta) + \log\binom{n}{y}\right)$$

EXPONENTIAL FAMILY CANONICAL PARAMETER $\varphi = \log\left(\frac{\theta}{1-\theta}\right)$.

IF Y FOLLOWS A GAMMA DISTRIBUTION WITH SCALE PARAMETER β AND SHAPE PARAMETER α, ITS pdf IS (α KNOWN)

$$f(y;\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\beta y}$$ dgamma FUNCTION IN R

$\Gamma(\cdot)$ IS THE gamma FUNCTION.

DISTRIBUTION BELONGS TO EXPONENTIAL FAMILY. WHY?

pdf CAN BE WRITTEN AS

$$f(y;\beta) = \exp\left[-\beta y + \alpha \log\beta - \log\Gamma(\alpha) + (\alpha-1)\log(y)\right]$$

SO $a(y)=y$, $b(\beta)=-\beta$, $c(\beta)=\alpha\log\beta-\log\Gamma(\alpha)$, $d(y)=(\alpha-1)\log(y)$

$E(Y)$ AND $VAR(Y)$?

COULD USE $E(a(Y)) = -c'(\beta)/b'(\beta)$ AND

$$VAR(a(Y)) = \frac{b''(\beta)c'(\beta) - c''(\beta)b'(\beta)}{[b'(\beta)]^3}$$ (EQUATIONS 3.9 & 3.12)

$c'(\beta)=\alpha/\beta$, $c''(\beta)=-\alpha/\beta^2$, $b'(\beta)=-1$, $b''(\beta)=0$

THEN,

$$E(Y) = \frac{-\alpha/\beta}{-1} = \frac{\alpha}{\beta}; \quad VAR(Y) = \frac{-(-\alpha/\beta^2)(-1)}{(-1)} = \frac{\alpha}{\beta^2}$$

$$f(y;\theta) = \exp[a(y)b(\theta)+c(\theta)+d(y)]$$ EXPONENTIAL FAMILY

$$\ell(\theta;y) = \log f(y;\theta) = a(y)b(\theta)+c(\theta)+d(y)$$

SCORE STATISTIC U

$$U(\theta;y) = \frac{d\,\ell(\theta;y)}{d\theta} = a(y)b'(\theta)+c'(\theta),$$ SINCE Y IS A RANDOM VARIABLE

$$U = a(Y)b'(\theta)+c'(\theta) \Rightarrow E(U) = b'(\theta)E(a(Y))+c'(\theta)$$

$$E(U) = b'(\theta)\left[\frac{-c'(\theta)}{b'(\theta)}\right] + c'(\theta) = -c'(\theta)+c'(\theta) = 0$$

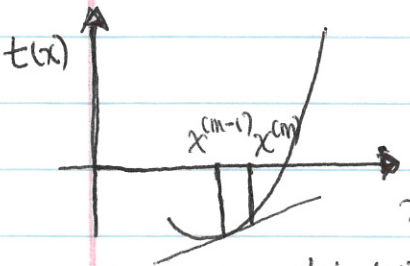
THE MEAN VALUE OF THE SCORE IS ZERO.

NOTES ON MAXIMUM LIKELIHOOD VIA THE NEWTON-RAPHSON METHOD

SUPPOSE WE HAVE A FUNCTION $t(x)$. DESCRIBES A SMOOTH CURVE

FIND SOLUTION TO $t(x) = 0$. ITERATING METHOD

CURRENT POINT $x^{(m-1)}$. HOW TO MOVE TO $x^{(m)}$?



SLOPE AT $x^{(m-1)}$, $t'(x^{(m-1)})$

IF $x^{(m-1)} \longrightarrow x^{(m)}$ IS A SMALL STEP

$$t'(x^{(m-1)}) \approx \frac{t(x^{(m)}) - t(x^{(m-1)})}{x^{(m)} - x^{(m-1)}} = \frac{\text{DIFF IN } t}{\text{STEP SIZE}}$$

WE WANT $t(x^{(m)}) = 0$. SOLUTION IS:

$$x^{(m)} = x^{(m-1)} - \frac{t(x^{(m-1)})}{t'(x^{(m-1)})}, \quad m = 1, 2, 3, \ldots$$

FOR NUMERICAL MAXIMUM LIKELIHOOD ESTIMATION

LIKELIHOOD FUNCTION $L(\theta; Y)$, θ PARAMETER Y DATA

$$\ell = \ell(\theta; y) = \log L(\theta; Y)$$

SCORE FUNCTION $U = \frac{d\ell}{d\theta}$. FIND SOLUTION TO $U(\theta) = 0$

SET AN INITIAL POINT $\theta^{(0)}$ (OR $\theta^{(1)}$).

UPDATE θ WITH, $\theta^{(m)} = \theta^{(m-1)} - \frac{U^{(m-1)}}{U'^{(m-1)}}$

$\theta^{(m-1)}$ IS THE CURRENT ITERATION. $\theta^{(m)}$ IS THE NEW POINT

$U^{(m-1)} = U(\theta^{(m-1)})$ SCORE FUNCTION AT $\theta^{(m-1)}$

$U'^{(m-1)}$ IS THE DERIVATIVE OF U AT $\theta^{(m-1)}$ OR

$$U'^{(m-1)} = \left. \frac{d^2\ell}{d\theta^2} \right|_{\theta = \theta^{(m-1)}}$$

COMMON TO WORK WITH $E(U')$ INSTEAD OF U' DIRECTLY.

FOR GLMs

$E(y_i) = \mu_i$, $g(\mu_i) = x_i^T \beta = \eta_i$, $x_i^T = (x_{i1}, x_{i2}, \ldots x_{ip})$

ESTIMATE THE VECTOR $\beta = (\beta_1, \ldots \beta_p)^T$ p-COEFFICIENTS.

NEWTON-RAPHSON LEADS INTO THE ITERATIVE EQUATION

$$X^T W X \underline{b}^{(m)} = X^T W \underline{z}$$

WHERE $X_{n \times p}$ MATRIX, COLS. ARE PREDICTOR VARIABLES.

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ \vdots & & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix}.$$

$\underline{b}^{(m)}$ ARE THE PARAMETER ESTIMATES FOR β.

W IS A $n \times n$ MATRIX, DIAGONAL, WITH ELEMENTS

$w_{ii} = \frac{1}{VAR(y_i)} \left(\frac{d\mu_i}{d\eta_i}\right)^2$, $i = 1, 2, \ldots n$ SO

$$W = \begin{pmatrix} w_{11} & 0 & \cdots & 0 \\ 0 & w_{22} & & 0 \\ & & & \vdots \\ 0 & \cdots & & w_{nn} \end{pmatrix},$$

IF $g(\mu_i) = \eta_i \Rightarrow \mu_i = g^{-1}(\eta_i)$

$\Rightarrow$ W MAY DEPEND ON b $(b^{(m-1)})$. THROUGH η_i.

NOTE THAT $X^T W X$ IS A $p \times p$ MATRIX.

FINALLY, $\underline{z}$ IS A VECTOR WITH ELEMENTS.

FOR $i = 1, 2. \; n$. $z_i = \sum_{k=1}^{p} x_{ik} b_k^{(m-1)} + (y_i - \mu_i)\left(\frac{d\eta_i}{d\mu_i}\right)$.

MOST STATISTICAL PACKAGES, WILL DO (ITERATIVE WEIGHTED L.S.)

1) INITIAL ESTIMATE FOR β, $b^{(0)}$.

2) COMPUTE $\underline{z}$, W AT $b^{(0)}$.

3) COMPUTE $b^{(1)} = (X^T W X)^{-1} X^T W \underline{z}$ OR EVEN BETTER SOLVE THE SYSTEM $X^T W X b^{(1)} = X^T W \underline{z}$

4) GO TO STEP (1) BUT REPLACE $b^{(0)}$ WITH $b^{(1)}$.

STOP WHEN $\|b^{(m)} - b^{(m-1)}\| < \epsilon \Rightarrow b^{(m)}$ IS THE MLE FOR β.

EXAMPLE: EXCERCISE 4.1 POISSON REGRESSION

Y_i = NO. OF AIDS CASES IN AUSTRALIA 1984-1988.

QUARTERLY DATA.

i = TIME PERIOD $i = 1, 2, \ldots 20$ (COVARIATE).

STATISTICAL MODEL $Y_i \sim \text{POISSON}(\lambda_i)$ WITH $\lambda_i = i^\theta$

OR $\log \lambda_i = \theta \log i$ - NO INTERCEPT. AS TIME PASSES MORE COUNTS ARE OBSERVED.

EXTEND GLM LOG-LINK FUNCTION TO

$\eta_i = g(\lambda_i) = \log \lambda_i = \beta_1 + \beta_2 \log i$ (INCLUDES INTERCEPT $\theta = \beta_2$)

$\beta = (\beta_1, \beta_2)^T$

$$\underset{20\times 2}{X} = \begin{bmatrix} 1 & \log(1) \\ 1 & \log(2) \\ \vdots & \vdots \\ 1 & \log(20) \end{bmatrix}$$

SINCE $\eta_i = \log \lambda_i \Rightarrow \frac{d\eta_i}{d\lambda_i} = \frac{1}{\lambda_i} = \frac{1}{e^{\beta_1 + \beta_2 \log i}}$

$\lambda_i = e^{\beta_1 + \beta_2 \log i}$

OR $\lambda_i = e^{\eta_i} \Rightarrow \frac{d\lambda_i}{d\eta_i} = e^{\eta_i} = e^{\beta_1 + \beta_2 \log i}$

IF $Y_i \sim \text{POISSON}(\lambda_i)$, $\lambda_i = \text{VAR}(Y_i) = e^{\beta_1 + \beta_2 \log i}$

$$w_{ii} = \frac{1}{e^{\beta_1 + \beta_2 \log i}} \left(e^{\beta_1 + \beta_2 \log i} \right)^2 = e^{\beta_0 + \beta_1 \log i}.$$

(FOLLOWS FROM $\frac{1}{\text{VAR}(Y_i)} \left(\frac{d\mu_i}{d\eta_i} \right)^2$).

$$z_i = \beta_1^{(m-1)} + \beta_2^{(m-1)} \log(i) + \Big(y_i - \underbrace{e^{\beta_1^{(m-1)} + \beta_2^{(m-1)} \log i}}_{\mu_i}\Big) \underbrace{\left(\frac{1}{e^{\beta_1^{(m-1)} + \beta_2^{(m-1)} \log(i)}} \right)}_{\frac{d\eta_i}{d\lambda_i}}$$

$$W = \begin{bmatrix} e^{\beta_1 + \beta_2 \log(1)} & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & & & \vdots \\ & & \ddots & & \\ 0 & \cdots & & & e^{\beta_1 + \beta_2 \log(20)} \end{bmatrix}_{20\times 20}.$$

$$\underset{2\times 20}{X^T} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \log(1) & \log(2) & \cdots & \log(20) \end{bmatrix}$$

$$X^T W = \begin{bmatrix} e^{\beta_1+\beta_2\log(1)} & e^{\beta_1+\beta_2\log(2)} & \cdots & e^{\beta_1+\beta_2\log(20)} \\ \log(1)\, e^{\beta_1+\beta_2\log(1)} & \cdots & \cdots & \log(20)\, e^{\beta_1+\beta_2\log(20)} \end{bmatrix}_{2\times 20}$$

$$X^T W X = \begin{bmatrix} \sum_{i=1}^{20} e^{\beta_1+\beta_2\log(i)} & \sum_{i=1}^{20} \log(i)\, e^{\beta_1+\beta_2\log(i)} \\ \sum_{i=1}^{20} \log(i)\, e^{\beta_1+\beta_2\log(i)} & \sum_{i=1}^{20} \log(i)^2\, e^{\beta_1+\beta_2\log(i)} \end{bmatrix}_{2\times 2}$$

$$X^T W \underline{z} = \begin{bmatrix} \sum_{i=1}^{20} z_i\, e^{\beta_1+\beta_2\log(i)} \\ \sum_{i=1}^{20} z_i \log(i)\, e^{\beta_1+\beta_2\log(i)} \end{bmatrix}$$

$X^T W X\, \underline{b}^{(1)} = X^T W \underline{z}$ LEADS TO TWO EQUATIONS:

$$\left[\sum_{i=1}^{20} e^{\beta_1^{(0)}+\beta_2^{(0)}\log(i)} \right] \beta_1^{(1)} + \left[\sum_{i=1}^{20} \log(i)\, e^{\beta_1^{(0)}+\beta_2^{(0)}\log(i)} \right] \beta_2^{(1)}$$

$$= \sum_{i=1}^{20} z_i\, e^{\beta_1^{(0)}+\beta_2^{(0)}\log(i)}.$$

$$\left[\sum_{i=1}^{20} \log(i)\, e^{\beta_1^{(0)}+\beta_2^{(0)}\log(i)} \right] \beta_1^{(1)} + \left[\sum_{i=1}^{20} [\log(i)]^2\, e^{\beta_1^{(0)}+\beta_2^{(0)}\log(i)} \right] \beta_2^{(1)}$$

$$= \sum_{i=1}^{20} z_i \log(i)\, e^{\beta_1+\beta_2\log(i)}.$$

CHECK R-CODE ON LINE.