REVIEW MATERIAL FOR STAT 474/574

$Y_1, Y_2, \ldots Y_n$ FOLLOW A POISSON (θ) DISTRIBUTION

$\theta = e^{\beta}$

FIND $\hat{\beta}$ THAT MINIMIZES $S = \sum_i (y_i - e^{\beta})^2$

$$\frac{dS}{d\beta} = 2\sum_i (y_i - e^{\beta})(-e^{\beta}) = 0 \Rightarrow \sum_i (y_i - e^{\beta}) = 0$$

SINCE $e^{\beta} > 0 \Rightarrow \sum_i y_i = ne^{\beta}$ $\quad \hat{\beta} = \log\left(\frac{\sum y_i}{n}\right)$

MLE OF θ

LIKELIHOOD FUNCTION: $L(\theta, y) = \dfrac{e^{-n\theta}\,\theta^{\sum y_i}}{\prod y_i!}$ OR

$$\ell(\theta, y) = -n\theta + (\textstyle\sum y_i)\log(\theta) - \log \prod y_i!$$

$$\frac{d\ell(\theta; y)}{d\theta} = -n + \frac{\sum_i y_i}{\theta} = 0 \Rightarrow \hat{\theta} = \bar{Y} \qquad \theta = e^{\beta} \Rightarrow \hat{\beta} = \log(\bar{Y}).$$

SUPPOSE $i = 1, 2, 3, \ldots 8$ REPRESENTS A GROUP $y_i \sim \text{POISSON}(\mu_i)$

$\mu_i = n_i e^{\theta \cdot i}$ WHERE n_i IS SOME POPULATION SIZE FOR GROUP i

SHOW THAT THIS IS A GLM.

$$f(y_i|\mu_i) = \frac{e^{-\mu_i}(\mu_i)^{y_i}}{y_i!} = e^{-\mu_i}\, e^{y_i \log(\mu_i)} - e^{\log(y_i!)}$$

$$= \exp\left[y_i \log(\mu_i) - \mu_i - \log(y_i!)\right]$$

BELONGS TO EXPONENTIAL FAMILY.

SINCE $\mu_i = n_i e^{\theta \cdot i} \Rightarrow \log(\mu_i) = \log(n_i) + \theta \cdot i$

LINK FUNCTION $g(\mu) = \log(\mu)$ $\quad x_i^t\beta = [\log(n_i), i]\begin{bmatrix} 1 \\ \theta \end{bmatrix}$

GROUP BINOMIAL DATA / LOGISTIC REGRESSION MODEL

$Y_i \sim \text{BINOMIAL}(n_i, \pi_i)$ $\qquad f(y_i | n_i, \pi_i) = \binom{n_i}{y_i} \pi_i^{y_i}(1-\pi_i)^{n_i - y_i}$

WITH ONE PREDICTOR X_i

$\log\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_0 + \beta_1 x_i$ so $g(\pi) = \log\left(\frac{\pi}{1-\pi}\right)$

WHAT IS THE LOG-LIKELIHOOD OF THE SATURATED MODEL?

$\beta = [\pi_1, \pi_2, \quad \pi_N]$

$$\log f(y_i | n_i, \pi_i) = \log\binom{n_i}{y_i} + y_i \log(\pi_i) + (n_i - y_i)\log(1-\pi_i)$$

$$\frac{d \log f(y_i | n_i, \pi_i)}{d\pi_i} = \frac{y_i}{\pi_i} - \frac{(n_i - y_i)}{(1-\pi_i)} = 0 \Rightarrow y_i - y_i\pi_i = n_i\pi_i - y_i\pi_i$$

$$\Rightarrow \hat{\pi}_i = y_i/n_i$$

COMBINE ALL OBSERVATIONS:

$$l(b_{max}; y) = \sum_i \log\binom{n_i}{y_i} + \sum_i y_i \log(y_i/n_i) + \sum_i (n_i - y_i)\log\left(1 - \frac{y_i}{n_i}\right)$$

EXCERSISE 7.2. 2×2 TABLE FOR EXPOSURE AND DISEASE

	DISEASED	NOT DISEASED
EXPOSED	π_1	$1-\pi_1$
NON-EXPOSED	π_2	$1-\pi_2$

ODDS OF DISEASE FOR EACH GROUP $O_i = \pi_i/(1-\pi_i)$, $i = 1,2$

ODDS RATIO

$$\phi = \frac{O_1}{O_2} = \frac{\pi_1(1-\pi_2)}{\pi_2(1-\pi_1)}.$$

(a) IF $\pi_i = e^{\beta_i}/(1+e^{\beta_i})$ AND $\beta_1 = \beta_2$ THEN $\phi = 1$

$\pi_1 = \pi_2 = e^{\beta}/(1+e^{\beta})$ $\quad (1-\pi_1) = (1-\pi_2) = 1/(1+e^{\beta})$

$$\phi = \frac{e^{\beta}/(1+e^{\beta})^2}{e^{\beta}/(1+e^{\beta})^2} = 1.$$

(b) J 2×2 TABLES ONE FOR EACH LEVEL x_j

$$\pi_{ij} = \frac{\exp(\alpha_i + \beta_i x_j)}{1+\exp(\alpha_i + \beta_i x_j)}; \quad i = 1,2,\ j = 1, \ldots J$$

SUPPOSE $\beta_1 = \beta_2$

ODDS RATIO FOR TABLE j IS

$$\phi = \frac{\pi_{1j}(1-\pi_{2j})}{\pi_{2j}(1-\pi_{1j})} = \frac{\exp(\alpha_1+\beta_1 x_j)}{\exp(\alpha_2+\beta_2 x_j)} = \exp(\alpha_1-\alpha_2+(\beta_1-\beta_2)x_j)$$

$$\pi_{1j} = \frac{\exp(\alpha_1+\beta_1 x_j)}{1+\exp(\alpha_1+\beta_1 x_j)}, \quad \pi_{2j} = \frac{\exp(\alpha_2+\beta_2 x_j)}{1+\exp(\alpha_2+\beta_2 x_j)}$$

$$1-\pi_{1j} = \frac{1}{(1+\exp(\alpha_1+\beta_1 x_j))} \qquad 1-\pi_{2j} = \frac{1}{(1+\exp(\alpha_2+\beta_2 x_j))}$$

IF $\beta_1=\beta_2 \Rightarrow \beta_1-\beta_2=0$ AND $\log(\phi)=\alpha_1-\alpha_2$

CONSTANT FOR j.

EXERCISE 8.1

FOR $J=2$ MODELS IN (8.4), (8.13), (8.15) AND (8.16)

REDUCE TO LOGISTIC REGRESSION.

IN (8.4)

$$\text{logit}(\pi_j) = \log\left(\frac{\pi_j}{\pi_1}\right) = x_j^T\beta_j \quad j=2=J$$

$$\Rightarrow \log\left(\frac{\pi_2}{\pi_1}\right) = x_j^T\beta_j$$ WITH 2 CATEGORIES $\pi_2 = 1-\pi_1$

$$\log\left(\frac{1-\pi_1}{\pi_1}\right) = x_j^T\beta_j$$

CUMULATIVE LOGIT MODEL

$$\log\left(\frac{\pi_1+\pi_2+\ldots+\pi_j}{\pi_{j+1}+\ldots+\pi_J}\right) = x_j^T\beta_j \qquad J=2$$

$$\log\left(\frac{\pi_1}{\pi_2}\right) = x^T\beta$$

ADJACENT CATEGORY MODEL: $\log\left(\frac{\pi_j}{\pi_{j+1}}\right) = \beta_{0j}+\beta_1 x_1+\ldots+\beta_{p-1}x_{p-1}$

$$\log\left(\frac{\pi_1}{\pi_2}\right) = \beta_{01}+\beta_1 x_1+\ldots\beta_{p-1}x_{p-1}$$

EXERCISE 9.1

LET $Y_1, Y_2, \dots Y_N$ BE INDEPENDENT $Y_i \sim Po(\mu_i)$

AND $\log \mu_i = \beta_1 + \sum_{j=2}^{J} x_{ij} \beta_j$, $i = 1, 2, \dots N$

(a) SCORE STATISTIC FOR β_1 : $\frac{d\,l(\beta; \underline{y})}{d\beta_1} = U_1$

$$l(\beta; \underline{y}) = -\sum_i \mu_i + \sum_i y_i \log(\mu_i) - \sum_i \log(y_i!)$$

$$\frac{d\,l(\beta; y)}{d\beta_1} = -\sum_i \frac{d\mu_i}{d\beta_1} + \sum_i y_i \left(\frac{d\mu_i}{d\beta_1}\right)\left(\frac{1}{\mu_i}\right)$$

SINCE $\log \mu_i = \beta_1 + \sum_{j=2}^{J} x_{ij} \beta_j \Rightarrow \mu_i = e^{\beta_1 + \sum_{j=2}^{J} x_{ij}\beta_j}$

$$\Rightarrow \frac{d\mu_i}{d\beta_1} = \mu_i \frac{d\beta_1}{d\beta_1} = \mu_i$$

$$\Rightarrow U_1 = -\sum_i \mu_i + \sum_i y_i (\mu_i)\left(\frac{1}{\mu_i}\right) = \sum_i (y_i - \mu_i)$$

(b) FOR MLE WE NEED $U_1 = 0 \Rightarrow \sum_i y_i = \sum_i \hat{\mu}_i$

(c) DEVIANCE IN 9.6 TAKES THE FORM

$$D = 2 \sum_i \left[o_i \log\left(\frac{o_i}{e_i}\right) - (o_i - e_i) \right] = \sum_i o_i \log\left(\frac{o_i}{e_i}\right) - \sum_i (o_i - e_i)$$

BUT IN OUR MODEL $o_i = y_i$ AND $e_i = \hat{\mu}_i \Rightarrow \sum_i (o_i - e_i) = 0$

AND

$$D = \sum_i o_i \log\left(\frac{o_i}{e_i}\right).$$

$Y_1, Y_2, \ldots Y_N$ ARE INDEPENDENT EXPONENTIAL(θ)

$f(y|\theta) = \theta e^{-\theta y}$

DESCRIBE THE NEWTON-RAPHSON METHOD TO ESTIMATE θ.

LIKELIHOOD FUNCTION

$$L(\theta; y) = \theta^n e^{-\theta \sum_{i=1}^{n} y_i}$$

$$\ell(\theta; y) = n \log(\theta) - \theta \sum_{i=1}^{n} y_i$$

RECURSION: $$\theta^{(m)} = \theta^{(m-1)} - \frac{U^{(m-1)}}{U'^{(m-1)}}$$

$$U = \frac{d\ell}{d\theta} = \frac{n}{\theta} - \sum_{i=1}^{n} y_i \quad , \quad U' = \frac{dU}{d\theta} = \frac{-n}{\theta^2}$$

FOR $m = 1, 2, \ldots$

$$\theta^{(m)} = \theta^{(m-1)} - \frac{\left(\frac{n}{\theta^{(m-1)}}\right) - \sum_{i=1}^{n} y_i}{\left(\frac{-n}{(\theta^{(m-1)})^2}\right)} = \theta^{(m-1)} + \left(\theta^{(m-1)} - [\theta^{(m-1)}]^2 \bar{y}\right)$$

WHAT IS THE HAZARD FUNCTION FOR THIS DISTRIBUTION?

HAZARD FUNCTION:

$$h(y) = \frac{f(y;\theta)}{S(y;\theta)} = \frac{f(y;\theta)}{1 - F(y;\theta)} = \frac{\theta e^{-\theta y}}{e^{-\theta y}} = \theta$$

SO PROPORTIONAL HAZARDS IS A MODEL OF THE FORM

$$\theta = \exp(\beta^t x) = \exp(\beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p)$$

REMARK:

$$F(y;\theta) = \int_0^y \theta e^{-\theta t} dt = -e^{-\theta t} \Big|_0^y = 1 - e^{-\theta y}$$

SUPPOSE WE HAVE THE DATA

x_i	-2	-1	0	1	2	PREDICTOR
y_i	1	2	1	3	5	RESPONSE

ASSUME $y_i \sim \text{POISSON}(\mu_i)$

(a) WRITE A GLM FOR EXPRESSING μ_i IN TERMS OF THE PREDICTOR x_i WITH AN IDENTITY LINK

$\mu_i = \beta_0 + \beta_1 x_i$

(b) REPEAT PART (a) BUT NOW USING A LOG-LINK

$\log(\mu_i) = \beta_0 + \beta_1 x_i$

(c) NOW EXPLAIN HOW YOU WOULD FIT MODELS (a) AND (b) USING R.

```
X = c(-2, -1, 0, 1, 2)
y = c(1, 2, 1, 3, 5)
# MODEL (a)
glm(y~x, family=poisson(link="identity"))
# MODEL (b)
glm(y~x, family=poisson(link="log"))
```

(d) HOW DO YOU DECIDE BETWEEN THE 2 MODELS?

log LINK ASSURES THAT $\mu_i > 0$.

EXTRACT AIC FOR BOTH MODELS ($aic IN R)

CHOOSE MODEL (b) IF

AIC (MODEL B) < AIC (MODEL A).

SINCE BOTH MODELS HAVE 2 PARAMETERS. COMPARE

D (MODEL A) WITH D(MODEL B)

WHERE D = DEVIANCE.

SUPPOSE WE HAVE A MODEL WHERE (TWO LEVEL)

$\mu_{jk} = E(Y_{jk}) = \mu + \alpha_j + \beta_k \quad j=1,2,\ K=1,2,3 \quad Y_{jk} \sim N(\mu_{jk}, \sigma^2)$

WRITE THE EQUATION FOR μ_{jk} IN MATRIX NOTATION $X\beta$

IS THIS A GLM?

$$Y = \begin{bmatrix} Y_{11} \\ Y_{12} \\ Y_{13} \\ Y_{21} \\ Y_{22} \\ Y_{23} \end{bmatrix}, \quad \underline{\mu} = \begin{bmatrix} \mu+\alpha_1+\beta_1 \\ \mu+\alpha_1+\beta_2 \\ \mu+\alpha_1+\beta_3 \\ \mu+\alpha_2+\beta_1 \\ \mu+\alpha_2+\beta_2 \\ \mu+\alpha_2+\beta_3 \end{bmatrix} \qquad \underline{\beta} = \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}$$

$$X = \begin{array}{c} \begin{matrix} \mu & \alpha_1 & \alpha_2 & \beta_1 & \beta_2 & \beta_3 \end{matrix} \\ \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \end{array} \Rightarrow \underline{\mu} = X\beta \quad \text{GLM WITH IDENTITY LINK.}$$

HOW DOES $X\beta$ CHANGE WITH THE RESTRICTIONS $\alpha_1+\alpha_2 = 1$

$\beta_1+\beta_2+\beta_3 = 1$

$\alpha_1 = -\alpha_2 \quad , \quad \beta_3 = -(\beta_2+\beta_3)$

$$\beta = \begin{bmatrix} \mu \\ \alpha_1 \\ \beta_1 \\ \beta_2 \end{bmatrix}, \qquad X = \begin{array}{c} \begin{matrix} \mu & \alpha_1 & \beta_1 & \beta_2 \end{matrix} \\ \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & 0 \\ 1 & -1 & 0 & 1 \\ 1 & -1 & -1 & -1 \end{bmatrix} \end{array} \qquad \underline{\mu} = X\beta .$$