3. Fisher : $A^{(t)}\beta^{(t+1)} = A^{(t)}\beta^{(t)} - u^{(t)}$

$$= -X^T W^{(t)} X \beta^{(t)} - X^T W^{(t)} \begin{pmatrix} (y_1-\mu_1^{(t)})\, g'(\mu_1^{(t)}) \\ \vdots \\ (y_i-\mu_i^{(t)})\, g'(\mu_i^{(t)}) \end{pmatrix}$$

Newton : $H^{(t)}\beta^{(t+1)} = H^{(t)}\beta^{(t)} - u^{(t)}$

$$= H^{(t)}\beta^{(t)} - X^T W^{(t)} \begin{pmatrix} (y_1-\mu_1^{(t)})\, g'(\mu_1^{(t)}) \\ (y_i-\mu_i^{(t)})\, g'(\mu_i^{(t)}) \end{pmatrix}$$

where $H_{rs} = \sum_{i=1}^{n} \left[\frac{y_i-\mu_i}{a_i(\phi)} x_{ir} \frac{\partial(\bar{\omega}_i g'(\mu_i))}{\partial \beta_s} + \frac{\bar{\omega}_i g'(\mu_i)}{a_i(\phi)} x_{ir} \frac{\partial(-\mu_i)}{\partial \beta_s} \right]$

* under canonical link $\eta_i = \theta_i$

Then:

$$\frac{1}{g'(\mu_i)} = \frac{1}{\left(\frac{\partial \eta_i}{\partial \mu_i}\right)} = \frac{\partial \mu_i}{\partial \eta_i} = \frac{\partial b'(\theta_i)}{\partial \theta_i} = b''(\theta_i)$$

and:

$$\bar{\omega}_i = \frac{1}{V(g'(\mu_i))^2}$$

so: $$\bar{\omega}_i\, g'(\mu_i) = \frac{1}{V(g'(\mu_i))^2} \cdot g'(\mu_i) = \frac{1}{V(g'(\mu_i))} = \frac{1}{V} \cdot \frac{\partial \mu_i}{\partial \eta_i} = \frac{1}{b''(\theta_i)} \cdot b''(\theta_i) = 1$$

Thus:

$$H_{rs} = \sum_{i=1}^{n} \left[\frac{y_i-\mu_i}{a_i(\phi)} x_{ir} \frac{\partial(1)}{\partial \beta_s} + \frac{\bar{\omega}_i}{a_i(\phi)} g'(\mu_i)\, x_{ir} \frac{\partial(-\mu_i)}{\partial \beta_s} \right]$$

$$= -\sum_{i=1}^{n} \left[\frac{\bar{\omega}_i}{a_i(\phi)} g'(\mu_i)\, x_{ir}\, x_{is} \right]$$

$$H = -X^T W X$$

Newton Raphson then becomes:

$$-X^T W X \beta^{(t)(t+1)} = -X^T W X^{(t)} \beta^{(t+1)} - u^{(t)}$$

which is identical to Fisher and $W = \begin{bmatrix} \frac{1}{a_1(\phi)} & & \\ & \ddots & \\ & & \frac{1}{a_n(\phi)} \end{bmatrix}$

4) fit 1

p.d.f Gaussian

Random: $Y_i \overset{indep}{\sim} N(\mu_i, \sigma^2)$ $i \in \{1, \ldots, 200\}$ $y_i \in \{0, 1 \ldots\}$

Systematic: $\eta_i = \beta_0 + \beta_1 x_i$

Link: $g(\mu_i) = \ln(\mu_i)$

$x_i \in \{1, \ldots 200\}$

fit 2

Random. $Y_i \overset{indep}{\sim} N(\mu_i, \sigma^2)$ $i \in \{1, \ldots 200\}$
$y_i \in \{1, \ldots 200\}$

Systematic: $\eta_i = \beta_0 + \beta_1 x_i$

Link: $g(\mu_i) = \mu_i$

$x \in \{1, \ldots 200\}$

fit 3

Random $Y_i \overset{indep}{\sim} Poisson(\mu_i)$ $i \in \{1, \ldots, 200\}$
$\mu_i > 0$ $y_i \{1, \ldots, 200\}$

Systematic: $\eta_i = \beta_0 + \beta_1 x_i$

Link $g(\mu_i) = \ln(\mu_i)$

$x_i \in \{1, \ldots, 200\}$

* Both fit 1 and fit 3 use a log link, while fit 2 uses an identity link. Fit 1 and Fit 2 are modeling the distribution of Y_i as gaussian, where fit 3 models it as a Poisson. Random compenent differs between fit 1 & 3, but links are same

5a $Y_i = Z_i/m_i$ where $Z_i \sim \text{Binom}(m_i, \pi_i)$

$f_{Y_i}(y_i) =$

$P(Y_i = y_i) = P(Z_i = m_i y_i) = \binom{m_i}{m_i y_i} \pi_i^{m_i y_i} (1-\pi_i)^{m_i - m_i y_i}$

$$= \exp\left[m_i y_i \ln(\pi) + (m_i - m_i y_i)\ln(1-\pi_i) + \ln\binom{m_i}{m_i y_i}\right]$$

$$= \exp\left[m_i y_i \ln(\pi) + m_i \ln(1-\pi_i) - m_i y_i \ln(1-\pi_i) + \ln\binom{m_i}{m_i y_i}\right]$$

$$= \exp\left[m_i y_i \ln(\pi_i) - m_i y_i (1-\pi_i) + m_i \ln(1-\pi_i) + \ln\binom{m_i}{m_i y_i}\right]$$

$$= \exp\left[m_i y_i \ln\left(\frac{\pi_i}{1-\pi_i}\right) + m_i \ln(1-\pi_i) + \ln\binom{m_i}{m_i y_i}\right]$$

$$= \exp\left[m_i \left(\frac{m_i y_i}{m_i} \text{logit}(\pi_i) + \ln(1-\pi_i)\right) + \ln\binom{m_i}{m_i y_i}\right)$$

$$= \exp\left[\frac{y_i \,\text{logit}(\pi_i) + \ln(1-\pi_i)}{1/m_i} + \ln\binom{m_i}{m_i y_i}\right]$$

$\theta_i = \text{logit}(\pi_i)$ $\qquad b(\theta_i) = -\ln(1-\pi_i)$

$\pi_i = \dfrac{e^{\theta_i}}{1+e^{\theta_i}}$ $\qquad = -\ln\left(1 - \dfrac{e^{\theta_i}}{1+e^{\theta_i}}\right)$

$\qquad = -\ln\left(\dfrac{1+e^{\theta_i} - e^{\theta_i}}{1+e^{\theta_i}}\right)$

$a(\phi) = \dfrac{1}{m_i}$

$c(y_i, \phi) = \ln\binom{m_i}{m_i y_i}$ $\qquad = -\ln\left(\dfrac{1}{1+e^{\theta_i}}\right)$ ← log rule $-\log_a b = \log_a\left(\frac{1}{b}\right)$

$\qquad = \ln(1+e^{\theta_i})$

$\dfrac{\partial b(\theta_i)}{\theta_i} = \dfrac{\partial(\ln(1+e^{\theta_i}))}{\theta_i}$ * chain rule

$= \dfrac{\partial(\ln(u))}{u} \cdot \dfrac{\partial(1+e^{\theta_i})}{\theta_i}$

$= \dfrac{1}{u} \cdot e^{\theta_i} = \dfrac{1}{1+e^{\theta_i}} \cdot e^{\theta_i} = \dfrac{e^{\theta_i}}{1+e^{\theta_i}} = \pi_i$

$\dfrac{b'(\theta_i)}{\theta_i} = \dfrac{\partial\left(\frac{e^{\theta_i}}{1+e^{\theta_i}}\right)}{\theta_i} =$ * quotient rule

→

$$\frac{\partial\left(\frac{e^{\theta_i}}{1+e^{\theta_i}}\right)}{\theta_i}$$

let $f(x) = e^{\theta_i}$

$g(x) = 1 + e^{\theta_i}$

<u>quotient rule</u>

$$\frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$$

$$= \frac{e^{\theta_i}(1+e^{\theta_i}) - (e^{\theta_i})^2}{(1+e^{\theta_i})^2}$$

$$= \frac{e^{\theta_i} + \cancel{(e^{\theta_i})^2} - \cancel{(e^{\theta_i})^2}}{(1+e^{\theta_i})^2}$$

$$= \frac{e^{\theta_i}}{(1+e^{\theta_i})^2}$$

$$= \frac{1}{(1+e^{\theta_i})} \cdot \frac{e^{\theta_i}}{1+e^{\theta_i}}$$

$$= \frac{1}{1+e^{\theta_i}} \cdot \pi_i$$

$$= \pi_i(1-\pi_i)$$

$$Var(Y_i) = a(\phi)\, b''(\theta_i)$$

$$= \frac{1}{m}\left(\pi_i(1-\pi_i)\right)$$

$$= \frac{\pi_i(1-\pi_i)}{m_i}$$

5b) $Y_i \sim N(\mu_i, \sigma^2)$

$$f_{Y_i}(y_i \mid \mu_i, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(y_i-\mu_i)^2}{2\sigma^2}\right]$$

$$= \exp\left\{ \ln\left[\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_i-\mu_i)^2}{2\sigma^2}\right)\right]\right\}$$

$$= \exp\left\{-\frac{1}{2}\ln(2\pi\sigma^2) - \frac{y_i + 2y_i\mu_i - \mu_i^2}{2\sigma^2}\right\}$$

$$= \exp\left\{-\frac{1}{2}\ln(2\pi\sigma^2) - \frac{y_i}{2\sigma^2} + \frac{2y_i\mu_i - \mu_i^2}{2\sigma^2}\right\}$$

$$= \exp\left\{\frac{2y_i\mu_i - \mu_i^2}{2\sigma^2} - \frac{1}{2}\ln(2\pi\sigma^2) - \frac{y_i}{2\sigma^2}\right\}$$

$$= \exp\left\{\frac{\cancel{2}y_i\mu_i}{\cancel{2}\sigma^2} - \frac{\mu_i^2}{2\sigma^2} - \frac{1}{2}\left(\ln(2\pi\sigma^2) + \frac{y_i}{\sigma^2}\right)\right\}$$

$$= \exp\left\{\frac{1}{\sigma^2}\left[y_i\mu_i - \frac{\mu_i^2}{2}\right] - \frac{1}{2}\left[\ln(2\pi\sigma^2) + \frac{y_i}{\sigma^2}\right]\right\}$$

$$= \exp\left\{\frac{y_i\mu_i - \mu_i^2/2}{\sigma^2} - \frac{1}{2}\left[\ln(2\pi\sigma^2) + \frac{y_i}{\sigma^2}\right]\right\}$$

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

$\qquad\qquad = \theta_i^2/2$

$a(\varnothing) = \sigma^2$

$\varnothing = \sigma^2$ $\quad c(y_i, \varnothing) = \ln(2\pi\sigma^2) + \frac{y_i}{\sigma^2}$

$\qquad\qquad = \ln(2\pi\varnothing) + \frac{y_i}{\varnothing}$

$\frac{\partial b(\theta_i)}{\theta_i} = \mu_i$

$\mathrm{Var}(Y_i) = a(\varnothing)\, b''(\theta_i)$

$\frac{\partial b'(\theta_i)}{\theta_i} = 1$

$\qquad = \sigma^2$

5c $Y_i \sim \text{Poisson}(\mu_i)$ if $Y_i \in \{0, 1, 2 \ldots\}$

$$f_{Y_i}(y_i \mid \mu_i) = \frac{\mu_i^{Y_i} e^{-\mu_i}}{Y_i!}$$

$$= \exp\left\{ \ln\left(\frac{\mu_i^{Y_i} e^{-\mu_i}}{Y_i!} \right) \right\}$$

$$= \exp\left\{ y_i \ln(\mu_i) - \mu_i - \ln(Y_i!) \right\}$$

$\theta_i = \ln(\mu_i)$ $\quad b(\theta_i) = \mu_i$

$\mu_i = e^{\theta_i}$ $\quad = e^{\theta_i}$

$a(\phi) = 1$ $\quad c(y_i, \phi) = -\ln(Y_i!)$

$\phi = 1$

$b'(\theta_i) = e^{\theta_i}$ $\quad Var(Y_i) = 1 \cdot e^{\theta_i}$

$b''(\theta_i) = e^{\theta_i}$ $\quad = e^{\theta_i}$

$= \mu_i$

6a)

Random

Let $Y_i = Z_i$ where $Z_i \overset{indep}{\sim} \text{Bernoulli}(\pi_i)$ $i = (1, \ldots, 32)$

$$f_{Y_i}(y_i) = P(Y_i = y_i) = P(Z_i = y_i) = \pi_i^{y_i}(1-\pi_i)^{1-\pi_i}$$

Systematic

$$\eta_i = \beta_1 hp_i + \sum_{k=2}^{4} \beta_k \mathbb{1}\{cyl_i = 2k\}$$

Link

$$g(\mu_i) = g(\pi_i) = \text{logit}(\pi_i) = \ln\left(\frac{\pi_i}{1-\pi_i}\right)$$

$Y_i = am_i =$ transimission type of ith car

where $am_i = 1$ is a manual transmission

and $am_i = 0$ is an automatic transmission

$$\mathbb{1}\{cyl_i = k\} = \begin{cases} 1 & cyl_i = k \\ 0 & cyl_i \neq k \end{cases}$$

$cyl_i =$ ith number of cyclinders $\in \{4, 6, 8\}$

$hp_i =$ horse power of ith car

ith = # of car observed

6b) + odds of automatic transmission:

$$= \frac{1}{e^{0.03797(hp) - 2.02874}} = \frac{1}{e^{0.03797(hp)} e^{-2.02874}}$$

+ odds of automatic transmission w/ one unit increase in horse power

$$= \frac{1}{e^{0.03797 + 0.03797(hp) - 2.02874}} = \frac{1}{e^{0.03797} e^{0.03797(hp)} e^{-2.02874}}$$

+ change in odds

$$= \frac{1}{e^{0.03797}} = 0.96$$

6c

$$\pi_i = \frac{e^{\beta_1 x_{i1} + \sum_{k=2}^{4} \beta_k 1(cyl_i = 2k)}}{1 + e^{\beta_1 x_{i1} + \sum_{k=2}^{4} \beta_k 1(cyl_i = 2k)}}$$

$$= \frac{e^{0.03797(100) - 2.02874}}{1 + e^{0.03797(100) - 2.02874}}$$

$$\pi_i = 0.854$$

6d) Random

Let $Y_i = Z_i$ where $Z_i \overset{indep}{\sim} \text{Bernoulli}(\pi_i)$

$i = (1, \ldots, 32)$

$$f_{Y_i}(y_i) = P(Y_i = y_i) = P(Z_i = y_i) = \pi_i^{\,y_i}(1-\pi_i)^{1-\pi_i}$$

$y \in \{0, 1\}$

Systematic

$$\eta_i = \beta_0 + \beta_1 hp_i + \beta_2 cyl_i$$

Link

$$g(\mu_i) = g(\pi_i) = \text{logit}(\pi_i) = \ln\left(\frac{\pi_i}{1-\pi_i}\right)$$

$Y_i = am_i$ = the transmission type of the i^{th} car, where $am_i = 1$ is a manual transmission and $am_i = 0$ is an automatic transmission.

β_0 = intercept, which is the mean transmission type when $hp_i = cyl_i = 0$

hp_i = horse power of i^{th} car

cyl_i = # of cylinders of i^{th} car $\in \{4, 6, 8\}$

$\hat{\pi}_i \quad \frac{e^{3.8322}}{1+e^{3.8322}}$

6e) Random

Let $Y_i = Z_i$ where $Z_i \overset{indep}{\sim}$ Bernoulli (π_i)

$y_i \in \{0, 1\}$ $\quad i = (1, \dots 32)$

Systematic

$$\eta_i = \beta_1 hp_i + \sum_{k=2}^{4} \beta_k 1\{cyl_i = 2k\}$$

Link

$g(\mu_i) = g(\pi_i) = \Phi^{-1}(\pi_i)$ = probit link

$Y_i = am_i$ = transmission type of the i^{th} car

$am_i = 0$ is automatic transmission

$am_i = 1$ is manual transmission

hp_i = horsepower of i^{th} car.

$1\{cyl_i = k\} = \begin{cases} 1 & cyl_i = k \\ 0 & cyl_i \neq k \end{cases}$

cyl_i = number of cylinders in i^{th} car $\in \{4, 6, 8\}$

* this model uses a probit link rather than a logit link function

6f) No, the canonical link is $logit(\pi_i) = \ln\left(\frac{\pi_i}{1-\pi_i}\right)$ although it is hard to choose between this link & a probit link since there is not enough info in the data. Both perform similarly when the probability of an event is neither very likely or unlikely.