Stat 7249: HW 1 solutions

2 ~~1.~~ a)

The likelihood function of exponential family is

$$f(y,\theta,\phi) = \exp\left\{\frac{y\theta - b(\theta)}{a(\phi)} + c(y,\theta)\right\}$$

Then the generating function can be derived as

$$\begin{aligned}
M_y(t) = E(e^{ty}) &= \int e^{ty}\exp\left\{\frac{y\theta - b(\theta)}{a(\phi)} + c(y,\phi)\right\}dy \\
&= \int \exp\left\{\frac{y(\theta + a(\phi)t) - b(\theta + a(\phi)t)}{a(\phi)} + c(y,\phi)\right\}\exp\left\{\frac{b(\theta + a(\phi)t) - b(\theta)}{a(\phi)}\right\}dy \\
&= \exp\left\{\frac{b(\theta + a(\phi)t) - b(\theta)}{a(\phi)}\right\}
\end{aligned}$$

We have

$$K_y(t) = \log M_y(t) = \frac{b(\theta + a(\phi)t) - b(\theta)}{a(\phi)}$$

b)

$$\begin{aligned}
K_y'(t) &= [b'[\theta + a(\phi)t]a(\phi)]/a(\phi) \\
K_y''(t) &= b''[\theta + a(\phi)t]a(\phi) \\
K_y'''(t) &= b'''[\theta + a(\phi)t]a^2(\phi)
\end{aligned}$$

Then

$$K_1 = K_y'(0) = b'(\theta), \quad ,K_2 = K_y''(0) = a(\phi)b''(\theta), \quad ,K_3 = K_y'''(0) = a^2(\phi)b'''(\theta)$$

Taking derivatives of the moment generating function with respect to t, we have

$$\begin{aligned}
M_y' &= e^{K_y}K_y' \\
M_y'' &= e^{K_y}K_y'' + (K_y')^2e^{K_y} \\
M_y''' &= e^{K_y(t)}K_y''' + K_y''(K_y')e^{K_y} + 2e^{K_y}K_y'K_y'' + (K_y')^3e^{K_y}
\end{aligned}$$

So,

$$\begin{aligned}
M_1 &= M_y'(0) = K_1 \\
M_2 &= M_y''(0) = K_1^2 + K_2 \\
M_3 &= M_y'''(0) = K_1^3 + 3K_1K_2 + K_3
\end{aligned}$$

3 ~~2.~~ a).

The Gamma function can be reparameterized as

$$\begin{aligned}
f(x) &= \frac{1}{\Gamma(\nu)}\left(\frac{\mu}{\nu}\right)^2 x^{\nu-1}e^{\{-\frac{\nu}{\mu}x\}} \\
&= \exp\left\{\frac{x\theta - b(\theta)}{a(\phi)} + c(x,\nu)\right\}
\end{aligned}$$

where $\theta = -\frac{1}{\mu}$, $b(\theta) = -\log(-\theta)$, $a(\phi) = \frac{1}{\nu}$ and $c(x,\nu) = (\nu-1)\log x + \nu\log\nu - \log\Gamma(\nu)$.

b).

$g(\mu) = -\frac{1}{\mu}$

c).

$$\begin{aligned}
E(x) &= K_1 = b'(\theta) = \frac{\partial}{\partial\theta}\log(-\theta) = \mu, \\
V(x) &= K_2 = \frac{\mu^2}{\nu}
\end{aligned}$$

[redacted]

4 ~~1.~~

$Y_i \sim Bin(m_i, \pi_i)$, $i = 1, cdots, n$. We have

$$f(y_i) = \exp\left(y_i \log \frac{\pi_i}{1-\pi_i} + m_i \log(1-\pi_i) + \log c_i\right)$$

so, $\theta_i = \log \frac{\pi_i}{1-\pi_i} + m_i \log(1-\pi_i)$ and $b(\theta_i) = -m_i \log(1-\pi_i) = m_i \log(1+e^{\theta_i})$
Likelihood equation is

$$\sum_i \frac{y_i - \mu_i}{var(y_i)} x_{ij} \frac{d\mu_i}{d\eta_i} = 0, \quad j = 1, \cdots, p$$

where $\mu_i = m_i\pi_i$, $var(y_i) = m_i\pi_i(1-\pi_i) = \mu_i(1-\frac{\mu_i}{m_i})$ and $\eta_i = \theta_i = \log \frac{\pi_i}{1-\pi_i} = \log \frac{\mu_i}{m_i-\mu_i}$.
Then

$$\frac{d\eta_i}{d\mu_i} = \frac{m_i}{\mu_i(m_i-\mu_i)}$$

We obtain

$$\sum_i \frac{y_i-\mu_i}{var(y_i)} x_{ij} \frac{d\mu_i}{d\eta_i} = \sum_i \frac{y_i-\mu_i}{\mu_i(1-\frac{\mu_i}{m_i})} x_{ij} \frac{m_i}{\mu_i(m_i-\mu_i)} = \sum_i (y_i-\mu_i)x_{ij} = 0$$

Since $\hat\beta \sim N(\beta, J^{-1})$, for the (j,k)th element of J, we have

$$\begin{aligned}
E(-\frac{\partial L(\beta)}{\partial\beta_j\partial\beta_k}) &= \sum_i E(\frac{\partial l_i}{\partial\beta_j}\cdot\frac{\partial l_i}{\partial\beta_k}) \\
&= \sum_i E(\sum_i \frac{y_i-\mu_i}{var(y_i)} x_{ij}\frac{d\mu_i}{d\eta_i}\cdot\sum_i \frac{y_i-\mu_i}{var(y_i)} x_{ik}\frac{d\mu_i}{d\eta_i}) \\
&= \sum_i (\frac{x_{ij}x_{ik}}{var^2(y_i)}(\frac{d\mu_i}{d\eta_i})^2 E(y_i-\mu_i)^2) = \sum_i (\frac{x_{ij}x_{ik}}{var(y_i)}(\frac{d\mu_i}{d\eta_i})^2) \\
&= \sum_i x_{ij}x_{ik}\frac{\mu_i(m_i-\mu_i)}{m_i}
\end{aligned}$$

Deviance$= 2\sum_i(y_i\tilde\theta_i - b(\tilde\theta_i)) - 2\sum_i(y_i\hat\theta_i - b(\hat\theta_i))$,
where $\theta_i = \log\frac{\mu_i}{m_i-\mu_i}$, $\hat\theta_i = \log\frac{\hat\mu_i}{m_i-\hat\mu_i}$, $\tilde\theta_i = \log\frac{y_i}{m_i-y_i}$, $b(\theta_i) = m_i\log(1+e^{\theta_i})$, $b(\hat\theta_i) = m_i\log(1+e^{\hat\theta_i})$
and $b(\tilde\theta_i) = m_i\log(1+e^{\tilde\theta_i})$
So,

$$\begin{aligned}
D &= 2\sum_i(y_i\log\frac{y_i}{m_i-y_i} - m_i\log\frac{m_i}{m_i-y_i}) - 2\sum_i(y_i\log\frac{\hat\mu_i}{m_i-\hat\mu_i} - m_i\log\frac{m_i}{m_i-\hat\mu_i}) \\
&= \sum_i(y_i\log\frac{y_i}{\hat\mu_i} + (m_i-y_i)\log\frac{m_i-y_i}{m_i-\hat\mu_i})
\end{aligned}$$

5 ~~2.~~

For canonical link, $\frac{d\mu_i}{d\eta_i} = \frac{d\mu_i}{d\theta_i} = \frac{db'(\theta_i)}{\theta_i} = b''(\theta_i)$. We have

$$\frac{\partial l_i}{\partial\beta_j} = \frac{y_i-\mu_i}{var(y_i)}x_{ij}\frac{d\mu_i}{d\eta_i} = \frac{y_i-\mu_i}{a(\phi)b''(\theta)}x_{ij}b''(\theta) = \frac{y_i-\mu_i}{a(\phi)}x_{ij}$$

$$\frac{\partial^2 l_i}{\partial\beta_j\partial\beta_k} = \frac{x_{ij}}{a(\phi)}\cdot\frac{d\mu_i}{d\beta_k} = -\frac{x_{ij}}{a(\phi)}\cdot\frac{d\mu_i}{d\eta_i}\frac{d\eta_i}{d\beta_k} = -\frac{x_{ij}x_{ik}}{a(\phi)}b''(\theta_i)$$

Hence,

$$\begin{aligned}
J_{jk} &= -E(\frac{\partial^2 L(\beta)}{\partial\beta_j\partial\beta_k}) = -\sum_i E(\frac{\partial^2 l_i}{\partial\beta_j\partial\beta_k}) \\
&= \sum_i E(\frac{x_{ij}x_{ik}}{a(\phi)}b''(\theta_i)) = \sum_i(\frac{x_{ij}x_{ik}}{a(\phi)}b''(\theta_i)) \\
&= \sum_i \frac{\partial^2 l_i}{\partial\beta_j\partial\beta_k} = \frac{\partial^2 L(\beta)}{\partial\beta_j\partial\beta_k} = H_{jk}
\end{aligned}$$

$j, k = 1, \cdots, p$.

From here, we can see that Newton-Ralphson and Fisher Scoring algorithms are identical.