1) a) $Y_{ij} \sim Binomial\,(m_{ij}, \pi_{ij})$

b) $f\,(Y_{ij} \mid m_{ij}, \pi_{ij}) = \binom{m_{ij}}{Y_{ij}} (\pi_{ij})^{Y_{ij}} (1-\pi_{ij})^{m_{ij}-Y_{ij}}$

$\Rightarrow \log(f(Y_{ij} \mid m_{ij}, \pi_{ij})) = \log\binom{m_{ij}}{Y_{ij}} + Y_{ij}\log(\pi_{ij}) + (m_{ij}-Y_{ij})\log(1-\pi_{ij})$

$= \log\binom{m_{ij}}{Y_{ij}} + Y_{ij}[\log(\pi_{ij}) - \log(1-\pi_{ij})] + m_{ij}\log(1-\pi_{ij})$

$= \log\binom{m_{ij}}{Y_{ij}} + Y_{ij}\log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) + m_{ij}\log\left(1 - \frac{e^{\theta_i-\theta_j}}{1+e^{\theta_i-\theta_j}}\right)$

$= \log\binom{m_{ij}}{Y_{ij}} + Y_{ij}(\theta_i-\theta_j) + m_{ij}\log\left(1 - \frac{e^{\theta_i-\theta_j}}{1+e^{\theta_i-\theta_j}}\right)$

(because $\log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \theta_i - \theta_j \Rightarrow \frac{\pi_{ij}}{1-\pi_{ij}} = e^{\theta_i-\theta_j}$

$\Rightarrow 1 + e^{\theta_i-\theta_j} = \frac{1}{1-\pi_{ij}} \Rightarrow \frac{e^{\theta_i-\theta_j}}{1+e^{\theta_i-\theta_j}} = \pi_{ij}$)

$\Rightarrow \log(f(Y_{ij} \mid m_{ij}, \pi_{ij})) = \log\binom{m_{ij}}{Y_{ij}} + Y_{ij}(\theta_i-\theta_j) + m_{ij}\log\left(\frac{1}{1+e^{\theta_i-\theta_j}}\right)$

$= \underbrace{\log\binom{m_{ij}}{Y_{ij}}} + Y_{ij}(\theta_i-\theta_j) - m_{ij}\underbrace{\log(1+e^{\theta_i-\theta_j})}$

$= h(Y_{ij}) + Y_{ij}(\theta_i-\theta_j) - m_{ij}K(\theta_i-\theta_j)$

Here $h(Y_{ij}) = \log\binom{m_{ij}}{Y_{ij}}$ and $K(t) = \log(1+e^t)$

HW4

Giang Vu

2/16/2021

Homework 4

1.

(c) The design matrix is 42x7 (top matrix) where each row is a possible outcome of a two-team match. For example the first row is 1 in first column entry, -1 in the second column entry, and 0 for the remaining 5 entries, which represents the match of team 1 vs team 2, and team 1 wins, team 2 loses, the remaining 5 teams don't compete in this match. This follows for the rest of the design matrix.

But because of the symmetry of this matrix, we can have a design that is only 21x7 like the bottom matrix

$$
\begin{vmatrix}
1 & -1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & -1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & -1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & -1 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & -1 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & -1 \\
0 & 1 & -1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & -1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & -1 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & -1 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & -1 \\
0 & 0 & 1 & -1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & -1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & -1 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & -1 \\
0 & 0 & 0 & 1 & -1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & -1 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & -1 \\
0 & 0 & 0 & 0 & 1 & -1 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & -1 \\
0 & 0 & 0 & 0 & 0 & 1 & -1 \\
-1 & 1 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 1 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 1 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 1 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 1 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & -1 & 1 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 1 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 1 & 0 & 0 \\
0 & -1 & 0 & 0 & 0 & 1 & 0 \\
0 & -1 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & -1 & 1 & 0 & 0 & 0 \\
0 & 0 & -1 & 0 & 1 & 0 & 0 \\
0 & 0 & -1 & 0 & 0 & 1 & 0 \\
0 & 0 & -1 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & -1 & 1 & 0 & 0 \\
0 & 0 & 0 & -1 & 0 & 1 & 0 \\
0 & 0 & 0 & -1 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & -1 & 1 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & -1 & 1
\end{vmatrix}
$$

$$\begin{vmatrix}
1 & -1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & -1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & -1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & -1 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & -1 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & -1 \\
0 & 1 & -1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & -1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & -1 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & -1 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & -1 \\
0 & 0 & 1 & -1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & -1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & -1 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & -1 \\
0 & 0 & 0 & 1 & -1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & -1 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & -1 \\
0 & 0 & 0 & 0 & 1 & -1 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & -1 \\
0 & 0 & 0 & 0 & 0 & 1 & -1
\end{vmatrix}$$

(d) If each team has the same home field advantage, we could think of it as an intercept for our model, which then will make our design matrix become 21x8, with a new first column full of 1's to represent the intercept. The new design matrix is displayed below.

1	1	-1	0	0	0	0	0
1	1	0	-1	0	0	0	0
1	1	0	0	-1	0	0	0
1	1	0	0	0	-1	0	0
1	1	0	0	0	0	-1	0
1	1	0	0	0	0	0	-1
1	0	1	-1	0	0	0	0
1	0	1	0	-1	0	0	0
1	0	1	0	0	-1	0	0
1	0	1	0	0	0	-1	0
1	0	1	0	0	0	0	-1
1	0	0	1	-1	0	0	0
1	0	0	1	0	-1	0	0
1	0	0	1	0	0	-1	0
1	0	0	1	0	0	0	-1
1	0	0	0	1	-1	0	0
1	0	0	0	1	0	-1	0
1	0	0	0	1	0	0	-1
1	0	0	0	0	1	-1	0
1	0	0	0	0	1	0	-1
1	0	0	0	0	0	1	-1

2.

(a) When $Y \sim$ Poisson(μ), the pf for Y with parameter $\mu > 0$ is

$$\begin{aligned} f(y, \mu) &= \frac{\mu^y}{y!} e^{-\mu} \\ &= e^{ylog(\mu)} e^{-\mu} \frac{1}{y!} \\ &= exp\{ylog(\mu) - \mu\} \frac{1}{y!} \end{aligned} \tag{1}$$

Here $\theta = log(\mu)$, $\phi = 1$, $K(\theta) = \mu$, and $h(y, \phi) = \frac{1}{y!}$

When $Y \sim$ N(μ, σ^2), the pdf for Y with parameters μ and σ is

$$\begin{aligned} f(y, \mu, \sigma) &= \frac{1}{\sigma\sqrt{2\pi}} exp\{-\frac{1}{2}(\frac{y-\mu}{\sigma})^2\} \\ &= exp\{\frac{-\frac{1}{2}y^2 - \frac{1}{2}\mu^2 + y\mu}{\sigma^2}\} \frac{1}{\sigma\sqrt{2\pi}} \\ &= exp\{\frac{y\mu - \frac{1}{2}\mu^2}{\sigma^2}\} exp\{-\frac{y^2}{2\sigma^2}\} \frac{1}{\sqrt{2\pi\sigma^2}} \end{aligned} \tag{2}$$

Here $\theta = \mu$, $\phi = \sigma^2$, $K(\theta) = \frac{1}{2}\mu^2$, and $h(y, \phi) = exp\{-\frac{y^2}{2\sigma^2}\} \frac{1}{\sqrt{2\pi\sigma^2}}$

2) b) $Y \sim f(y, \theta, \phi) = \exp\left\{\frac{y\theta - K(\theta)}{\phi}\right\} h(y, \phi)$

The first derivative w.r.t θ is

$$f'(y, \theta, \phi) = \frac{y - K'(\theta)}{\phi} \exp\left\{\frac{y\theta - K(\theta)}{\phi}\right\} h(y, \phi)$$

$$\Rightarrow \int f'(y, \theta, \phi)\, dy = \int \frac{y - K'(\theta)}{\phi} f(y, \theta, \phi)\, dy$$

$$\Rightarrow \frac{\partial}{\partial\theta} \underbrace{\int f(y, \theta, \phi)\, dy}_{\text{a proper density}} = \frac{1}{\phi}\left(\int y f(y, \theta, \phi)\, dy - K'(\theta) \int f(y, \theta, \phi)\, dy\right)$$

$$\Rightarrow \frac{\partial}{\partial\theta}(1) = \frac{1}{\phi}\left(E(Y) - K'(\theta).(1)\right)$$

$$\Rightarrow 0 = \frac{E(Y) - K'(\theta)}{\phi}$$

$$\Rightarrow E(Y) = K'(\theta)$$

The second derivative w.r.t θ is

$$f''(y, \theta, \phi) = \left[\left(\frac{y - K'(\theta)}{\phi}\right)^2 - \frac{K''(\theta)}{\phi}\right] f(y, \theta, \phi)$$

$$\Rightarrow \int f''(y, \theta, \phi)\, dy = \int \left[\left(\frac{y - K'(\theta)}{\phi}\right)^2 - \frac{K''(\theta)}{\phi}\right] f(y, \theta, \phi)\, dy$$

$$\Rightarrow \frac{\partial^2}{\partial\theta^2} \int f(y, \theta, \phi)\, dy = \int \left(\frac{y - K'(\theta)}{\phi}\right)^2 f(y, \theta, \phi)\, dy - \frac{K''(\theta)}{\phi} \int f(y, \theta, \phi)\, dy$$

$$\Rightarrow 0 = \frac{1}{\phi^2} \int (y - E(Y))^2 f(y, \theta, \phi)\, dy - \frac{K''(\theta)}{\phi}$$

$$\Rightarrow 0 = \frac{1}{\phi^2} Var(Y) - \frac{K''(\theta)}{\phi}$$

$$\Rightarrow Var(Y) = \phi K''(\theta)$$

c) $Y_i \sim f(y; x_i^T\beta, 1) = \exp\left\{\frac{y x_i^T\beta - K(x_i^T\beta)}{1}\right\} h(y, 1)$

$E_{\hat\beta}(Y_i) = K'(x_i^T\hat\beta)$

$f'(y; x_i^T\beta, 1) = \frac{y - K'(x_i^T\beta)}{1} \exp\left\{\frac{y x_i^T\beta - K(x_i^T\beta)}{1}\right\} h(y, 1)$

Set $f'(y, x_i^T\beta, 1) = 0$, the MLE $\hat\beta$ will be the value that satisfies $y - K'(x_i^T\hat\beta) = 0$

(because $\exp\left\{\frac{y x_i^T\beta - K(x_i^T\beta)}{1}\right\} h(y, 1) = f(y, x_i^T\beta, 1)$

Since this is a well defined density, $f(y, x_i^T\beta, 1) > 0$

$\Rightarrow$ for $f'(y, x_i^T\beta, 1) = 0$, $y - K'(x_i^T\hat\beta) = 0$)

$\Rightarrow \quad y_i = K'(x_i^T\hat\beta) = E_{\hat\beta}(Y_i)$

$\Rightarrow \quad Y = E_{\hat\beta}(Y)$

$\Rightarrow \quad X^T\{Y - E_{\hat\beta}(Y)\} = X^T\{Y - Y\} = 0_p$

(Check second derivative of $f(y, x_i^T\beta, 1)$ evaluated at $\hat\beta$

$$f''(y; x_i^T\hat\beta, 1) = \left[\underbrace{\left(\frac{y_i - K'(x_i^T\hat\beta)}{1}\right)^2}_{0} - \underbrace{\frac{K''(x_i^T\hat\beta)}{1}}_{<0}\right] \underbrace{f(y; x_i^T\hat\beta, 1)}_{>0}$$

$\Rightarrow f''(y, x_i^T\hat\beta, 1) < 0$

$\Rightarrow \hat\beta$ is MLE of β)

2) d) Asymptotically,

$$\hat{\beta} \overset{n/p \text{ large}}{\sim} N(\beta, \underbrace{I_n^{-1}(\beta)}_{\text{Inverse of Fisher information matrix}})$$

From class, we learned that for logistic regression

$$I_n(\beta) = X^T Var(Y) X$$

From part (b), $Var(Y) = \phi K''(\theta)$

$\Rightarrow$ Asymptotic variance of $\hat{\beta}$ is

$$(X^T \phi K''(\theta) X)^{-1}$$