STAT 2132
Orly Olbum

HOMEWORK 4

1. $\pi_{ij} = P(\text{competitor } i \text{ beats } j)$

$$\log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \theta_i - \theta_j, \quad i \neq j, \quad i,j = 1, \ldots, K$$

say $K=7$ teams

m_{ij} = #games i & j play together

Y_{ij} = # games i beats j

ⓐ what is the distribution of Y_{ij} in terms of m_{ij} & π_{ij}?

$\pi_{ij} = P(Y_{ij}=1)$

Y_{ij} has binomial distribution, with $n = m_{ij}$ and $p = \pi_{ij}$

Binomial dist. takes the form:

$$f(x|n,p) = \binom{n}{x} p^x (1-p)^{n-x}$$

here, $n = m_{ij}$, $p = \pi_{ij} = Pr(Y_{ij}=1)$.

ⓑ Let $l_{ij}(\theta_i, \theta_j; Y_{ij})$ be log-likelihood for Θ. Show that

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

for some functions $K(t)$, $h(t)$.

Find an expression for $K(t)$.

If i & j teams play n games (m_{ij}), say $Y_{ij}=1$ if i wins & $Y_{ij}=0$ if i loses

$$L\, f(Y_{ij}|m_{ij}, \pi_{ij}) = \log\left[\binom{m_{ij}}{Y_{ij}} \pi_{ij}^{Y_{ij}} (1-\pi_{ij})^{m_{ij}-Y_{ij}}\right]$$

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

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

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

$$= \underbrace{\log\binom{m_{ij}}{Y_{ij}}}_{h(Y_{ij})} + \underbrace{Y_{ij}(\theta_i - \theta_j)}_{Y_{ij}(\theta_i-\theta_j)} \underbrace{- m_{ij}\log[1+\exp(\theta_i-\theta_j)]}_{m_{ij}K(\theta_i-\theta_j)}$$

$$K(t) = \log(1+e^t)$$

1, continued

ⓒ write 21 x 7 design matrix X for BT model.

7 teams $i=1, \dots, 7$, $j=1, \dots, 7$

for team i winning, Y=1

for team i winning, j loses, =-1

(see next page for matrix)

ⓓ Now suppose we have home field advantage taken into account. what is new X?

If team i is the home team, our model now has an intercept, say α, so that the model becomes

$$\log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \alpha + \theta_i - \theta_j$$

The new design matrix has a column of $\underline{1}$ at the first column.

match	team	1	2	3	4	5	6	7
1	2	1	-1	0	0	0	0	0
1	3	1	0	-1	0	0	0	0
1	4	1	0	0	-1	0	0	0
1	5	1	0	0	0	-1	0	0
1	6	1	0	0	0	0	-1	0
1	7	1	0	0	0	0	0	-1
2	3	0	1	-1	0	0	0	0
2	4	0	1	0	-1	0	0	0
2	5	0	1	0	0	-1	0	0
2	6	0	1	0	0	0	-1	0
2	7	0	1	0	0	0	0	-1
3	4	0	0	1	-1	0	0	0
3	5	0	0	1	0	-1	0	0
3	6	0	0	1	0	0	-1	0
3	7	0	0	1	0	0	0	-1
4	5	0	0	0	1	-1	0	0
4	6	0	0	0	1	0	-1	0
4	7	0	0	0	1	0	0	-1
5	6	0	0	0	0	1	-1	0
5	7	0	0	0	0	1	0	-1
6	7	0	0	0	0	0	1	-1

this matrix is actually 42x7 but there is a pattern based on who plays who

2. consider

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

$\phi \rightarrow$ dispersion parameter

$\theta \rightarrow$ canonical parameter

(a) what are θ, ϕ, and $K(\theta)$ if $Y \sim \text{Poisson}(\mu)$, $\mu > 0$?
what about if $Y \sim N(\mu, \sigma^2)$?

If $Y \sim \text{Poisson}(\mu)$, $f(y|\mu) = \frac{\mu^y e^{-\mu}}{y!}$

$$= \frac{1}{y!} \exp\{y \log \mu - \mu\}$$

which gives us

$$\theta = \log \mu$$
$$K(\theta) = \mu = \exp(\theta)$$
$$\phi = \frac{1}{y!} \qquad , \mu > 0$$

If $Y \sim N(\mu, \sigma^2)$, $f(y|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{-\frac{1}{2\sigma^2}(y-\mu)^2\right\}$

$$= \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{\frac{\mu}{\sigma^2} y - \frac{1}{2\sigma^2} y^2 - \frac{1}{2\sigma^2}\mu^2 - \log\sigma\right\}$$

$$\theta = \mu$$
$$K(\theta) = \frac{\theta^2}{2} = \frac{\mu^2}{2}$$
$$\phi = \sigma^2$$

(b) $Y \sim f(y;\theta,\phi)$. Find expressions for $E(Y)$, $var(Y)$ in terms of $K(\theta)$ and ϕ.

Take derivative:

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

Integrate both sides

$$\int f'(y;\theta,\phi)\,dy = \frac{\partial}{\partial\theta}\int f(y;\theta,\phi)\,dy = \frac{\partial}{\partial\theta}(1) = 0$$

(assuming we can interchange der. & integral)

$$\rightarrow \int \frac{y - K'(\theta)}{\phi} f(y;\theta,\phi)\,dy = \frac{1}{\phi}\left[\int y f(y;\theta,\phi)\,dy - K'(\theta)\int f(y;\theta,\phi)\,dy\right]$$

$$= \frac{1}{\phi}\left[E(Y) - K'(\theta)\right] = 0$$

so, $E(Y) = K'(\theta)$ $\rightarrow$

2, continued

var(Y):

take the second derivative of both sides

$$f''(y;\theta,\phi) = [(\frac{y-K'(\theta)}{\phi}) - \frac{K''(\theta)}{\phi}]\, f(y;\theta,\phi)$$

Integrate both sides

$$\int f''(y;\theta,\phi)\,dy = \frac{\partial^2}{\partial\theta^2}\int f(y;\theta,\phi)\,dy = \frac{\partial^2}{\partial\theta^2}(1) = 0$$

$$\rightarrow \int [\underbrace{(\frac{y-K'(\theta)}{\phi}}_{E(Y)^2 - E(Y^2) = var(Y)})^2 - \frac{K''(\theta)}{\phi}]\, f(y;\theta,\phi)\,dy$$

$$= \frac{1}{\phi}[\frac{var(Y)}{\phi} - K''(\theta)]$$

so, $var(Y) = \phi K''(\theta)$

ⓒ $Y_i \sim f(y; x_i^T\beta, \phi)$, $\beta \in \mathbb{R}^p$, $\phi > 0$ are unknown

If $Y_1, \dots, Y_n$ $\perp$, show that $\hat\beta$ satisfies

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

where $E_{\hat\beta}(Y)$ is $E(Y_i)$ for $Y_i \sim f(y; x_i^T\hat\beta, 1)$

If $y \sim f(y; x_i^T\beta, 1)$,

$$y \Rightarrow \exp\{y x_i^T\beta - K(x_i^T\beta)\}\, h(y,1)$$

from ⓑ

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

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

set $f' = 0$, solve for β to find $\hat\beta$

when $y_i = K'(x_i^T\hat\beta)$,

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

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

ⓓ what is asymptotic variance of $\hat\beta$ (inverse of Fisher information) in terms of K, β, X, ϕ?

From ⓒ and based on what we've done in class, $\hat\beta \to \beta$ is asymptotically normally dist'd

Fisher: $-E\{\nabla^2_\beta\, l(x_i^T\beta, \phi)\} = -E_{\hat\beta}\{\frac{\phi K''(\theta)}{\beta}\}$

so $var(\hat\beta) = \dfrac{1}{-E_{\hat\beta}\{\frac{\phi K''(\theta)}{\beta}\}}$