Stat 742 **Homework 6 Solutions** **Fall 2008**

6.3) X does not need to be full rank. The solution that will maximize the likelihood will not be unique in this case. Since the normal equation is still solved that it is still possible to obtain a MLE, which is non-unique.

6.4) MSE = $E[SSE/c - \sigma^2]^2 = \frac{1}{c^2}E[SSE^2] - \frac{1}{c}\, 2\sigma^2 E[SSE] + \sigma^4$

$E[SSE] = E[y'(I - P_x)y] = \sigma^2(N - r)$

$E[SSE^2] = V(SSE) + E(SSE)^2 = 2\sigma^4(N - r) + \sigma^4(N - r)^2$

So MSE = $\left(\frac{1}{c^2}\right)[2\sigma^4(N - r) + \sigma^4(N - r)^2] - \left(\frac{1}{c}\right)2\sigma^4(N - r) + \sigma^4$

$\frac{\partial MSE}{\partial c} = -2c^{-3}[2\sigma^4(N - r) + \sigma^4(N - r)^2] - c^{-2}2\sigma^4(N - r)^2 = 0$

$\Rightarrow 2 + (N - r) - c = 0$

$\Rightarrow c = N - r + 2$

$\frac{\partial^2 MSE}{\partial c^2} = 6c^{-4}[2\sigma^4(N - r) + \sigma^4(N - r)^2] + 2c^{-3}(N - r)^2 > 0$ at $c = N - r + 2$, so it's a min.

Thus, SSE/c is minimized when $c = N - r + 2$.

6.6) a) The density for e_i is $f(e_i) = 1/(2\sigma)\, I(-\sigma < e_i < \sigma)$

$\Rightarrow f(y_i) = 1/(2\sigma)\, I(b'x_i - \sigma < y_i < b'x_i + \sigma)$

So $L(b, \sigma|y) = [1/(2\sigma)]^n\, I(b'x_1 - \sigma < y_i < b'x_1 + \sigma) \ldots I(b'x_n - \sigma < y_n < b'x_n + \sigma)$

$L = [1/(2\sigma)]^n$ if for all y_i, $b'x_i - \sigma < y_i < b'x_i + \sigma$ and 0 otherwise.

So the MLE is any b in the set $\{b: -\sigma < y_i - b'x_i < \sigma \text{ for all } i=1,\ldots,n\}$

b) $f(y_i) = \frac{e^{b'x_i - y_i}}{(1+e^{b'x_i - y_i})^2}$

$L(b, \sigma|y) = \frac{\prod e^{b'x_i - y_i}}{\prod(1+e^{b'x_i - y_i})^2} \Rightarrow \log L = \sum(b'x_i - y_i) - 2\sum \log(1 + e^{b'x_i - y_i})$

So the MLE is the value of b minimizes $\sum b'x_i - 2\sum \log(1 + e^{b'x_i - y_i})$.

c) $f(y_i) = \exp\{-|y_i - b'x_i|/\sigma\}/(2\sigma)$

So $L(b, \sigma|y) = \exp\{-\Sigma|y_i - b'x_i|/\sigma\}/(2\sigma)^n$

$\log L = -\Sigma|y_i - b'x_i|/\sigma - n\log(2\sigma)$

The MLE is the value of b that minimizes $\Sigma|y_i - b'x_i|$. There is no close form solution to this problem. One can use linear programming to find its solution.

6.7) From exercise 4.1, $y_i = n_i p + \varepsilon_i$ and $y_i|p \sim \mathrm{Binom}(n_i, p)$

$$f(y) = \prod \binom{n_i}{y_i} p^{\Sigma y_i}(1-p)^{\Sigma(n_i - y_i)}$$

$$\log f(y) = \Sigma \log\binom{n_i}{y_i} + \Sigma y_i \log(p) + \Sigma(n_i - y_i)\log(1-p)$$

$$\frac{d\log f}{dp} = \frac{\Sigma y_i}{p} - \frac{\Sigma(n_i - y_i)}{1-p} = 0 \Rightarrow \hat{p}_{MLE} = \frac{\Sigma y_i}{\Sigma n_i} = \frac{\Sigma y_i}{N}$$

$$\frac{d^2\log f}{dp^2} = -\frac{\Sigma y_i}{p^2} - \frac{\Sigma(n_i - y_i)}{(1-p)^2} < 0, \text{ so it's a maximum.}$$

The MLE estimate is the same as the BLUE estimate.

6.9) In the model $y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk}$, $i=1,\dots,a$, $j=1,\dots,n$, $k=1,\dots,n_{ij}$

$$\mathbf{Xb} = \left[\begin{array}{c|ccc|c|c} & \mathbf{1_n} & & \mathbf{0} & \mathbf{I_n} & \\ \mathbf{1_{an}} & & \mathbf{1_n} & & \vdots & \mathbf{I_{an}} \\ & \mathbf{0} & & \mathbf{1_n} & \mathbf{I_n} & \end{array}\right]\begin{bmatrix} \mu \\ \boldsymbol{\alpha} \\ \boldsymbol{\beta} \\ \boldsymbol{\gamma} \end{bmatrix}, \text{where} \quad \begin{array}{r} \boldsymbol{\alpha} = [\alpha_1 \dots \alpha_a]' \\ \boldsymbol{\beta} = [\beta_1 \dots \beta_n]' \\ \boldsymbol{\gamma} = [\gamma_{11} \cdots \gamma_{1n}\ \gamma_{21} \cdots \gamma_{2n} \cdots \cdots \gamma_{a1} \cdots \gamma_{an}]' \end{array}$$

Without the interaction terms in the model, rank($\mathbf{X}$) = a + n – 1.

Notice that in the columns corresponding to each γ_{ij} for a fixed value of i, they sum up to the column vector corresponding to α_i, such that $\sum_{j=1}^{n} \gamma_{ij} = \alpha_i$.

Likewise, for the columns corresponding to each γ_{ij} for a fixed value of j, the columns sum up to the column vector corresponding to β_j, such that $\sum_{i=1}^{a} \gamma_{ij} = \beta_j$.

So $\sum_{j=1}^{n} \gamma_{ij}$ is confounded with α_i and $\sum_{i=1}^{a} \gamma_{ij}$ is confounded with β_j.

Thus, we have a + n non-estimable functions, but they are not yet linearly independent.

Note that the columns for α_i's, as well as the columns for β_j's, sum up to the first column.

Thus any a + n – 1 of the non-estimable functions $\sum_{j=1}^{n} \gamma_{ij}$ and $\sum_{i=1}^{a} \gamma_{ij}$ will be linearly independent and confounded with the (a + n – 1) linearly independent main effects.