STATS 3001 / STATS 4101 / STATS 7054
Statistical Modelling III
Tutorial 5
2022
Solutions

QUESTIONS:

1. Suppose $y_1, y_2, \ldots, y_n$ are independent Poisson observations, $y_i \sim Po(\mu_i)$ and consider the **linear** regression model

$$\boldsymbol{\mu} = X\boldsymbol{\beta}$$

where X is an $n \times p$ matrix with linearly independent columns and $\boldsymbol{\beta} \in \mathbb{R}^p$ is the vector of unknown parameters.

(a) Write down the log-likelihood

$$\ell(\boldsymbol{\mu}; \boldsymbol{y})$$

(b) Calculate $\dfrac{\partial \ell}{\partial \mu_i}$.

(c) Find the matrix of partial derivatives

$$\left[\frac{\partial \mu_i}{\partial \beta_j}\right].$$

(d) Show that the score vector can be expressed as

$$\mathcal{S}(\boldsymbol{\beta}) = \left[\frac{\partial \ell}{\partial \beta_j}\right] = X^T D_\mu^{-1}(\boldsymbol{y} - \boldsymbol{\mu})$$

where $D_\mu = diag(\mu_1, \mu_2, \ldots, \mu_n)$.

(e) Evaluate the Fisher information matrix.

(f) Describe the Fisher scoring algorithm.

i. State the iterative step of the algorithm to obtain $\hat{\boldsymbol{\beta}}^{(t+1)}$ from $\hat{\boldsymbol{\beta}}^{(t)}$.

ii. Suggest a starting value, $\hat{\boldsymbol{\beta}}^{(0)}$.

SOLUTIONS:

(a)

$$
\begin{aligned}
\ell(\boldsymbol{\mu};\boldsymbol{y}) &= \log \prod_{i=1}^{n} \frac{e^{-\mu_i}\mu_i^{y_i}}{y_i!} \\
&= \sum_{i=1}^{n}(y_i \log \mu_i - \mu_i) - \log \prod_{i=1}^{n} y_i!
\end{aligned}
$$

(b)

$$
\frac{\partial \ell}{\partial \mu_i} = \frac{y_i}{\mu_i} - 1.
$$

(c)

Observe first that

$$
\boldsymbol{\mu} = X\boldsymbol{\beta} \Rightarrow \mu_i = \sum_{k=1}^{p} x_{ik}\beta_k
$$

so that

$$
\frac{\partial \mu_i}{\partial \beta_j} = \frac{\partial}{\partial \beta_j}\sum_{k=1}^{p} x_{ik}\beta_k = \sum_{k=1}^{p}\frac{\partial}{\partial \beta_j} x_{ik}\beta_k = x_{ij}.
$$

Hence, we obtain in matrix notation

$$
\left[\frac{\partial \mu_i}{\partial \beta_j}\right] = X.
$$

(d)

Using the chain rule of multivariable calculus,

$$
\begin{aligned}
\frac{\partial \ell}{\partial \beta_j} &= \sum_{i=1}^{n} \frac{\partial \ell}{\partial \mu_i}\frac{\partial \mu_i}{\partial \beta_j} \\
&= \sum_{i=1}^{n}\left(\frac{y_i}{\mu_i} - 1\right) x_{ij} \\
&= \sum_{i=1}^{n}\frac{(y_i - \mu_i)}{\mu_i} x_{ij}
\end{aligned}
$$

which is the j^{th} element of

$$X^T D_\mu^{-1}(\boldsymbol{y} - \boldsymbol{\mu})$$

(e)

$$\mathcal{S}(\boldsymbol{\beta}) = X^T(\boldsymbol{Y} - \boldsymbol{\mu}) \Rightarrow Var(\mathcal{S}(\boldsymbol{\beta})) = X^T D_\mu^{-1} Var(Y) D_\mu^{-1} X$$

and since the Y_i are independent Poisson we have

$$Var(\boldsymbol{Y}) = diag(\mu_1, \mu_2, \ldots, \mu_n) = D_\mu.$$

Hence,

$$\mathcal{I}(\boldsymbol{\beta}) = X^T D_\mu^{-1} X.$$

(f)

$$\hat{\boldsymbol{\beta}}^{(t+1)} = \hat{\boldsymbol{\beta}}^{(t)} + \mathcal{I}(\hat{\boldsymbol{\beta}}^{(t)})^{-1}\mathcal{S}(\hat{\boldsymbol{\beta}}^{(t)}) = \hat{\boldsymbol{\beta}}^{(t)} + (X^T D_{\hat{\mu}^{(t)}}^{-1} X)^{-1} X^T D_{\hat{\mu}^{(t)}}^{-1}(\boldsymbol{y} - \hat{\boldsymbol{\mu}}^{(t)}).$$

(g)

$$\hat{\boldsymbol{\beta}}_0 = (X^T X)^{-1} X^T \boldsymbol{y}$$
