Information Matrix Equality

Step 1

Define the score of the log likelihood as

$$s_i(\theta) \equiv \nabla_\theta l_i(\theta)' = [\frac{\partial l_i}{\partial \theta_1}, ..., \frac{\partial l_i}{\partial \theta_P}]'$$

We now first show that

$$E[s_i(\theta)|x_i] = 0$$

This is simply to apply the definition of conditional expectation operator and the fact that $l_i(\theta) = log f(y_i|x_i;\theta)$

$$\begin{aligned} E[s_i(\theta)|x_i] &= \int s_i(\theta) f(y|x_i;\theta)\nu(dy) \\ &= \int \nabla_\theta f(y|x_i;\theta)\nu(dy) \\ &= \nabla_\theta \int f(y|x_i;\theta)\nu(dy) \\ &= 0 \end{aligned}$$

Step 2

Now we are equipped to show the conditional information matrix equality. Use the fact that $E[s_i(\theta)|x_i] = 0$ and take derivative on both side.

$$\begin{aligned} \nabla_\theta E[s_i(\theta)|x_i] &\equiv \nabla_\theta [\int s_i(\theta) f(y|x_i;\theta)\nu(dy)] \\ &= \int \nabla_\theta [s_i(\theta) f(y|x_i;\theta)]\nu(dy) = 0 \end{aligned}$$

Recall that $H_i(\theta) \equiv \nabla_\theta s_i(\theta) = \nabla_\theta^2 l_i(\theta)$. Thus

$$\nabla_\theta [s_i(\theta) f(y|x_i;\theta)] = H_i(\theta) f(y|x_i;\theta) + s_i(\theta)s_i(\theta)' f(y|x_i;\theta)$$

So we have

$$\begin{aligned} \int [H_i(\theta) + s_i(\theta)s_i(\theta)'] f(y|x_i;\theta)\nu(dy) = 0 \\ -E[H_i(\theta)|x_i] = E[s_i(\theta)s_i(\theta)'|x_i] \end{aligned}$$

Step 3

Now we evaluate both conditional expectations at θ_0, we reach the conditional information matrix equality

$$-E[H_i(\theta_0)|x_i] = E[s_i(\theta_0)s_i(\theta_0)'|x_i]$$

Use the law of iterated expectations (with respect to the distribution of x_i), we have $A_0 = B_0$.