We will use that $\underset{\nearrow}{I(\beta)} = E\Big(\underbrace{-\frac{\partial^2 l}{\partial\beta\partial\beta^T}}_{\substack{H(\beta)\\ |\\ \text{observed}\\ \text{Fisher info}}}\Big) = \underbrace{E\Big(\Big(\frac{\partial l}{\partial\beta}\Big)\Big(\frac{\partial l}{\partial\beta^T}\Big)\Big)}_{Cov(s(\beta))}$

expected Fisher info

Why can we do that?

---

1) Casella & Berger, Lemma 7.3.11 page 338
(StatInf textbook) (univariate case)

If $f(x|\theta)$ satisfies

$$\frac{d}{d\theta} E_\theta\left(\frac{\partial}{\partial\theta}\ln f(X|\theta)\right) = \int \frac{\partial}{\partial\theta}\left[\left(\frac{\partial}{\partial\theta}\ln f(x|\theta)\right)f(x|\theta)\right]dx$$

$$\| $$

$$\frac{d}{d\theta}\int\left(\frac{\partial}{\partial\theta}\ln f(x|\theta)\right)f(x|\theta)\,dx$$

[this is true for univ. exp. fam]

then:

$$E_\theta\left(\left(\frac{\partial}{\partial\theta}\ln f(X|\theta)\right)^2\right) = -E_\theta\left(\frac{\partial^2}{\partial\theta^2}\ln f(X|\theta)\right)$$

$$E_\theta\left(\frac{\partial^2}{\partial\theta^2}\ln f(X|\theta)\right) = \int \underbrace{\frac{\partial^2}{\partial\theta^2}(\ln f(x|\theta))}\,f(x|\theta)\,dx$$

$(uv)' = u\cdot v' + u'\cdot v$

$$\frac{\partial}{\partial\theta}\Big[\underbrace{\frac{1}{f(x|\theta)}}_{u}\cdot\underbrace{\frac{\partial}{\partial\theta}f(x|\theta)}_{v}\Big]$$

$$= \int\left[-\left(\frac{1}{f(x|\theta)}\right)^2\cdot\frac{\partial}{\partial\theta}f(x|\theta)\,\frac{\partial}{\partial\theta}f(x|\theta) + \frac{\partial^2}{\partial\theta^2}f(x|\theta)\cdot\frac{1}{f(x|\theta)}\right]f(x|\theta)\,dx$$

since $\quad \frac{\partial}{\partial\theta}\ln f(x|\theta) = \frac{1}{f(x|\theta)}\cdot\frac{\partial}{\partial\theta}f(x|\theta)$

$$= -\int \left[\frac{1}{f(x|\theta)} \cdot \frac{\partial}{\partial\theta} f(x|\theta)\right]^2 f(x|\theta)\, dx \;+\; \underbrace{\int \frac{\partial^2}{\partial\theta^2} f(x|\theta) \cdot dx}_{0}$$

0 from E, Var of exp-fam

$$= -E\left(\frac{\partial}{\partial\theta} \ln f(x)|\theta)\right)$$

qed

2) This can be generalized to the p-dimensional case. We will not prove that.

3) We still miss the connection between $Cov(s(\beta))$ and $E(s(\beta)s(\beta)^T)$

$$E(s(\theta)) = E\Big(\underbrace{\frac{\partial}{\partial\theta} \ln f(x|\theta)}_{\text{score}}\Big) =$$

$$\int \frac{\partial}{\partial\theta}(\ln f(x|\theta))\, f(x|\theta)\, dx = \int \frac{1}{f} \frac{\partial f}{\partial\theta} f \, dx$$

$$= \int \frac{\partial f}{\partial\theta}\, dx = 0 \leftarrow \text{see 1 proof for } E(Y) \text{ and } Var(Y)$$

Thus $E(s(\beta)) = 0$, so that

$$Cov(s(\beta)) = E\big((s(\beta) - 0)(s(\beta) - 0)^T\big) = E(s(\beta)\, s(\beta)^T)$$