Chapter 4

Lecture 13

The score function, Fisher information and bounds

Let Θ be an open interval in $\mathbb{R}^1$ and suppose that $dP_\theta(s) = \ell_\theta(s)d\mu(s)$, where μ is a fixed measure on S. Suppose that $\theta \mapsto \ell_\theta(s)$ is differentiable for each fixed s; then $\delta \mapsto \Omega_{\delta,\theta}(s) = \frac{\ell_\delta(s)}{\ell_\theta(s)}$ is also differentiable for each fixed (s,θ). If we use dashes for derivatives with respect to the parameters as described, then

$$\Omega'_{\theta,\theta}(s) = \frac{\ell'_\theta(s)}{\ell_\theta(s)} =: \gamma_\theta^{(1)}(s)$$

is the SCORE FUNCTION at θ (given s). We also define $I(\theta) := E_\theta\big(\gamma_\theta^{(1)}(s)\big)^2$, the FISHER INFORMATION (for estimating θ) in s.

Note.

$$\Big(\int_S \ell_\delta(s)d\mu(s) = 1 \ \forall\delta \in \Theta\Big)$$

$$\Rightarrow \Big(\int_S \Omega'_{\delta,\theta}(s)dP_\theta(s) = \int_S \frac{\ell'_\delta(s)}{\ell_\theta(s)}\ell_\theta(s)d\mu(s) = \int_S \ell'_\delta(s)d\mu(s) = 0 \ \forall\delta \in \Theta\Big)$$

$$\Rightarrow E_\theta\big(\gamma_\theta^{(1)}(s)\big) = E_\theta\big(\Omega'_{\theta,\theta}(s)\big) = 0 \Rightarrow I(\theta) = \operatorname{Var}_\theta(\gamma_\theta^{(1)})$$

Similarly, we have $\int_S \ell''_\delta(s)d\mu(s) = 0$, $\int_S \ell'''_\delta(s)d\mu(s) = 0$, etc. for all $\delta \in \Theta$, so that $E_\theta\big(\gamma_\theta^{(j)}(s)\big) = 0$ for $j = 1,2,3,\ldots$, where $\gamma_\theta^{(j)}(s) = \big(\frac{\partial^j \ell_\theta(s)}{\partial\theta^j}\big)/\ell_\theta(s)$. Conditions under which the interchanging of differentiation and integration (as above) is valid will be given later.

Suppose that we are interested in W_θ and want some concrete method of constructing it. We have that

$$\Omega_{\delta,\theta}(s) = \Omega_{\theta,\theta} + (\delta-\theta)\gamma_\theta^{(1)}(s) + \frac{1}{2}(\delta-\theta)^2\gamma_\theta^{(2)}(s) + \cdots,$$

which suggests that $W_\theta = \operatorname{Span}\{1, \gamma_\theta^{(1)}, \gamma_\theta^{(2)}, \ldots\}$. We will see that this equality holds exactly in a one-parameter exponential family and approximately in general in large