Math 281A Homework 6 Solution

1. Let $\{X_i, Y_i\}_{i=1}^n$ be i.i.d. random vectors with $Y_i \in \{0,1\}$, and

$$\mathbb{P}_{\alpha,\beta}(Y_i = 1 | X_i = x) = \frac{1}{1+e^{-\alpha-\beta x}}.$$

The distribution of X_i is non-degenerate, but unknown. Do we have closed form of MLE $(\hat{\alpha}, \hat{\beta})$? Derive the asymptotic distribution of $(\hat{\alpha}, \hat{\beta})$.

Solution: Denote

$$\Psi(u) = \frac{1}{1+e^{-u}},$$

then the log likelihood function is

$$\ell(\alpha,\beta) = \sum_{i=1}^n [Y_i \log(\Psi(\alpha+\beta X_i)) + (1-Y_i)\log(1-\Psi(\alpha+\beta X_i))],$$

and

$$\nabla \ell(\alpha,\beta) = \sum_{i=1}^n \frac{(Y_i - \Psi(\alpha+\beta X_i))\Psi'(\alpha+\beta X_i)}{\Psi(\alpha+\beta X_i)(1-\Psi(\alpha+\beta X_i))}\begin{bmatrix} 1 \\ X_i \end{bmatrix}.$$

The MLE $(\hat{\alpha}, \hat{\beta})$ is the root of the above display, which doesn't have closed form. For the asymptotic distribution of $(\hat{\alpha}, \hat{\beta})$, we calculate the Fisher information,

$$I_{\alpha,\beta} = \mathbb{E}\frac{(\Psi'(\alpha+\beta X))^2}{\Psi(\alpha+\beta X)(1-\Psi(\alpha+\beta X))}\begin{bmatrix} 1 & X \\ X & X^2 \end{bmatrix},$$

which is invertible when the distribution of X is non-degenerate. Hence,

$$\sqrt{n}\left(\begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} - \begin{bmatrix} \alpha \\ \beta \end{bmatrix}\right) \xrightarrow{d} N\left(0, I_{\alpha,\beta}^{-1}\right).$$

2. Let $\{X_i\}_{i=1}^n$ be i.i.d. from Poisson$(1/\theta)$.

   (a) Calculate the Fisher information I_θ in one observation;

   Solution: Log likelihood function is

   $$\ell(\theta) = -\frac{1}{\theta} - x\log\theta - \log x!.$$

   It can be calculated that

   $$I_\theta = \frac{1}{\theta^3}.$$

   (b) Derive the MLE $\hat{\theta}$ and show its asymptotic distribution.

   Solution: The MLE is

   $$\hat{\theta} = \frac{1}{\bar{X}},$$

   and we have

   $$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, \theta^3).$$

3. Let $\{X_i\}_{i=1}^n$ be i.i.d. from $N(\theta, \theta)$.

(a) Calculate the Fisher information I_θ in one observation;

Solution: Log likelihood function is

$$\ell(\theta) = -\frac{x^2}{2\theta} - \frac{\theta}{2} - \frac{1}{2}\log\theta + x - \frac{1}{2}\log(2\pi),$$

and the Fisher information is

$$I_\theta = \frac{2\theta + 1}{2\theta^2}.$$

(b) Derive the MLE $\hat{\theta}$ and show its asymptotic distribution.

Solution: The MLE is

$$\hat{\theta} = \frac{\sqrt{4\overline{X^2} + 1} - 1}{2},$$

and we have

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N\left(0, \frac{2\theta^2}{2\theta + 1}\right).$$

4. (a) Calculate the Kullback-Leibler divergence between two exponential distributions with different scale parameters, when is it maximal?

Solution: The Kullback-Leibler divergence is

$$M(\lambda) = \mathbb{E}_{\lambda_0} \log \frac{f_\lambda(X)}{f_{\lambda_0}(X)} = 1 - \frac{\lambda}{\lambda_0} + \log\frac{\lambda}{\lambda_0}.$$

It's maximal when $\lambda = \lambda_0$.

(b) Calculate the Kullback-Leibler divergence between two normal distributions with different location and scale parameters, when is it maximal?

Solution: The Kullback-Leibler divergence is

$$M(\mu, \sigma) = \mathbb{E}_{\mu_0,\sigma_0} \log \frac{f_{\mu,\sigma}(X)}{f_{\mu_0,\sigma_0}(X)} = \frac{1}{2} - \frac{(\mu_0 - \mu)^2 + \sigma_0^2}{2\sigma^2} + \log\frac{\sigma_0}{\sigma}.$$

It's maximal when $\mu = \mu_0$ and $\sigma = \sigma_0$.