Statistical Inference II, 2013, HW6

Random vectors $\mathbf{X}_i = \begin{bmatrix} X_{i1} \\ X_{i2} \end{bmatrix}, \quad i = 1,...,n$ follow (iid) bivariate normal distribution

$$f_{\boldsymbol{\mu},\boldsymbol{\Sigma}}(\mathbf{x}) = \frac{1}{2\pi |\boldsymbol{\Sigma}|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right\},$$

where $\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$ and $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11}^2 & \sigma_{12} \\ \sigma_{12} & \sigma_{22}^2 \end{bmatrix}$, and the parameter $\boldsymbol{\theta} = (\mu_1, \mu_2, \sigma_{11}^2, \sigma_{22}^2, \sigma_{12})^T$ is unknown.

1. Show that the score function is written as

$$\mathbf{S}_n(\boldsymbol{\theta}) = \sum_{i=1}^{n} \mathbf{U}_i(\boldsymbol{\theta}),$$

   where $\mathbf{U}_i(\boldsymbol{\theta}) = (U_{i1}(\boldsymbol{\theta}),..., U_{i5}(\boldsymbol{\theta}))^T$ has zero mean (show).

2. Derive the Fisher information matrix $I_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = nI_1(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ (explicitly).
3. Find $c(\boldsymbol{\theta}) = \Pr(X_{i1} \le X_{i2})$.
4. Find the asymptotically pivotal quantity (p.495) for $c(\boldsymbol{\theta})$ and then derive the $(1-\alpha)$-asymptotic confidence interval for $c(\boldsymbol{\theta})$.

Answer:

1. Let $|\boldsymbol{\Sigma}| = \sigma_{11}^2\sigma_{22}^2 - \sigma_{12}^2$ and

$$D_i(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{\left\{\sigma_{22}^2(x_{i1}-\mu_1)^2 - 2\sigma_{12}(x_{i1}-\mu_1)(x_{i2}-\mu_2) + \sigma_{11}^2(x_{i2}-\mu_2)^2\right\}}{|\boldsymbol{\Sigma}|} = (\mathbf{x}_i - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i - \boldsymbol{\mu}).$$

The log-likelihood is

$$\begin{aligned} l(\boldsymbol{\mu}, \boldsymbol{\Sigma}) &= -n\log(2\pi) - \frac{n}{2}\log|\boldsymbol{\Sigma}| - \frac{1}{2}\sum_{i=1}^{n}(\mathbf{x}_i - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i - \boldsymbol{\mu}) \\ &= -n\log(2\pi) - \frac{n}{2}\log|\boldsymbol{\Sigma}| - \frac{1}{2}\sum_{i=1}^{n}D_i(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \end{aligned}$$

Note that

$$\frac{\partial D_i(\boldsymbol{\mu}, \boldsymbol{\Sigma})}{\partial \sigma_{11}^2} = \frac{1}{|\boldsymbol{\Sigma}|}\left\{(x_{i2}-\mu_2)^2 - \sigma_{22}^2 D_i(\boldsymbol{\mu}, \boldsymbol{\Sigma})\right\}, \quad E\left(\frac{\partial D_i(\boldsymbol{\mu}, \boldsymbol{\Sigma})}{\partial \sigma_{11}^2}\right) = -\frac{\sigma_{22}^2}{|\boldsymbol{\Sigma}|}$$

$$\frac{\partial D_i(\boldsymbol{\mu}, \boldsymbol{\Sigma})}{\partial \sigma_{22}^2} = \frac{1}{|\boldsymbol{\Sigma}|}\left\{(x_{i1}-\mu_1)^2 - \sigma_{11}^2 D_i(\boldsymbol{\mu}, \boldsymbol{\Sigma})\right\}, \quad E\left(\frac{\partial D_i(\boldsymbol{\mu}, \boldsymbol{\Sigma})}{\partial \sigma_{22}^2}\right) = -\frac{\sigma_{11}^2}{|\boldsymbol{\Sigma}|},$$

$$\frac{\partial D_i(\boldsymbol{\mu}, \boldsymbol{\Sigma})}{\partial \sigma_{12}} = \frac{2}{|\boldsymbol{\Sigma}|}\left\{-(x_{i1}-\mu_1)(x_{i2}-\mu_2) + \sigma_{12} D_i(\boldsymbol{\mu}, \boldsymbol{\Sigma})\right\}, \quad E\left(\frac{\partial D_i(\boldsymbol{\mu}, \boldsymbol{\Sigma})}{\partial \sigma_{12}}\right) = \frac{2\sigma_{12}}{|\boldsymbol{\Sigma}|},$$

and hence

$$\frac{\partial l(\boldsymbol{\mu}, \boldsymbol{\Sigma})}{\partial \mu_1} = \sum_{i=1}^{n}\frac{1}{|\boldsymbol{\Sigma}|}\{\sigma_{22}^2(x_{i1}-\mu_1) - \sigma_{12}(x_{i2}-\mu_2)\} = \sum_{i=1}^{n}U_{i1}(\boldsymbol{\theta}),$$

$$\frac{\partial l(\boldsymbol{\mu}, \boldsymbol{\Sigma})}{\partial \mu_2} = \sum_{i=1}^{n}\frac{1}{|\boldsymbol{\Sigma}|}\{\sigma_{11}^2(x_{i2}-\mu_2) - \sigma_{12}(x_{i1}-\mu_1)\} = \sum_{i=1}^{n}U_{i2}(\boldsymbol{\theta}),$$

$$\frac{\partial l(\boldsymbol{\mu}, \boldsymbol{\Sigma})}{\partial \sigma_{11}^2} = \sum_{i=1}^{n}\frac{1}{|\boldsymbol{\Sigma}|}\left\{-\frac{\sigma_{22}^2}{2} - \frac{(x_{i2}-\mu_2)^2}{2} + \frac{\sigma_{22}^2}{2}D_i(\boldsymbol{\mu}, \boldsymbol{\Sigma})\right\} = \sum_{i=1}^{n}U_{i3}(\boldsymbol{\theta}),$$

$$\frac{\partial l(\boldsymbol{\mu}, \boldsymbol{\Sigma})}{\partial \sigma_{22}^2} = \sum_{i=1}^{n}\frac{1}{|\boldsymbol{\Sigma}|}\left\{-\frac{\sigma_{11}^2}{2} - \frac{(x_{i1}-\mu_1)^2}{2} + \frac{\sigma_{11}^2}{2}D_i(\boldsymbol{\mu}, \boldsymbol{\Sigma})\right\} = \sum_{i=1}^{n}U_{i4}(\boldsymbol{\theta}),$$

$$\frac{\partial l(\boldsymbol{\mu}, \boldsymbol{\Sigma})}{\partial \sigma_{12}} = \sum_{i=1}^{n}\frac{1}{|\boldsymbol{\Sigma}|}\left\{\sigma_{12} + (x_{i1}-\mu_1)(x_{i2}-\mu_2) - \sigma_{12}D_i(\boldsymbol{\mu}, \boldsymbol{\Sigma})\right\} = \sum_{i=1}^{n}U_{i5}(\boldsymbol{\theta}).$$

2. Note that

$$-E\left(\frac{\partial^2 l(\boldsymbol{\mu}, \boldsymbol{\Sigma})}{\partial \mu_1^2}\right) = n\frac{\sigma_{22}^2}{|\boldsymbol{\Sigma}|}, \quad -E\left(\frac{\partial^2 l(\boldsymbol{\mu}, \boldsymbol{\Sigma})}{\partial \mu_2^2}\right) = n\frac{\sigma_{11}^2}{|\boldsymbol{\Sigma}|}, \quad -E\left(\frac{\partial^2 l(\boldsymbol{\mu}, \boldsymbol{\Sigma})}{\partial \mu_2 \partial \mu_1}\right) = -n\frac{\sigma_{12}}{|\boldsymbol{\Sigma}|},$$

$$-E\left(\frac{\partial^2 l(\boldsymbol{\mu},\boldsymbol{\Sigma})}{\partial\sigma_{11}^2\partial\mu_1}\right)=-E\left(\frac{\partial^2 l(\boldsymbol{\mu},\boldsymbol{\Sigma})}{\partial\sigma_{22}^2\partial\mu_1}\right)=-E\left(\frac{\partial^2 l(\boldsymbol{\mu},\boldsymbol{\Sigma})}{\partial\sigma_{12}\partial\mu_1}\right)=0\,,$$

$$-E\left(\frac{\partial^2 l(\boldsymbol{\mu},\boldsymbol{\Sigma})}{\partial\sigma_{11}^2\partial\mu_2}\right)=-E\left(\frac{\partial^2 l(\boldsymbol{\mu},\boldsymbol{\Sigma})}{\partial\sigma_{22}^2\partial\mu_2}\right)=-E\left(\frac{\partial^2 l(\boldsymbol{\mu},\boldsymbol{\Sigma})}{\partial\sigma_{12}\partial\mu_2}\right)=0\,,$$

$$-E\left(\frac{\partial^2 l(\boldsymbol{\mu},\boldsymbol{\Sigma})}{\partial(\sigma_{11}^2)^2}\right)=n\frac{\sigma_{22}^4}{2|\boldsymbol{\Sigma}|^2},\quad -E\left(\frac{\partial^2 l(\boldsymbol{\mu},\boldsymbol{\Sigma})}{\partial(\sigma_{22}^2)^2}\right)=n\frac{\sigma_{11}^4}{2|\boldsymbol{\Sigma}|^2},\quad -E\left(\frac{\partial^2 l(\boldsymbol{\mu},\boldsymbol{\Sigma})}{\partial(\sigma_{12})^2}\right)=n\frac{2(\sigma_{11}^2\sigma_{22}^2+\sigma_{12}^2)}{2|\boldsymbol{\Sigma}|^2}$$

$$-E\left(\frac{\partial^2 l(\boldsymbol{\mu},\boldsymbol{\Sigma})}{\partial\sigma_{11}^2\partial\sigma_{22}^2}\right)=n\frac{\sigma_{12}^2}{2|\boldsymbol{\Sigma}|^2},\quad -E\left(\frac{\partial^2 l(\boldsymbol{\mu},\boldsymbol{\Sigma})}{\partial\sigma_{11}^2\partial\sigma_{12}}\right)=n\frac{\sigma_{12}\sigma_{22}^2}{2|\boldsymbol{\Sigma}|^2},\quad -E\left(\frac{\partial^2 l(\boldsymbol{\mu},\boldsymbol{\Sigma})}{\partial\sigma_{22}^2\partial\sigma_{12}}\right)=n\frac{\sigma_{12}\sigma_{11}^2}{2|\boldsymbol{\Sigma}|^2}\,.$$

Thus, the Fisher information matrix is

$$I_1(\boldsymbol{\mu},\boldsymbol{\Sigma})=\begin{bmatrix}\frac{\sigma_{22}^2}{|\boldsymbol{\Sigma}|} & -\frac{\sigma_{12}}{|\boldsymbol{\Sigma}|} & 0 & 0 & 0\\ -\frac{\sigma_{12}}{|\boldsymbol{\Sigma}|} & \frac{\sigma_{11}^2}{|\boldsymbol{\Sigma}|} & 0 & 0 & 0\\ 0 & 0 & \frac{\sigma_{22}^4}{2|\boldsymbol{\Sigma}|^2} & \frac{\sigma_{12}^2}{2|\boldsymbol{\Sigma}|^2} & -\frac{\sigma_{12}\sigma_{22}^2}{|\boldsymbol{\Sigma}|^2}\\ 0 & 0 & \frac{\sigma_{12}^2}{2|\boldsymbol{\Sigma}|^2} & \frac{\sigma_{11}^4}{2|\boldsymbol{\Sigma}|^2} & -\frac{\sigma_{12}\sigma_{11}^2}{|\boldsymbol{\Sigma}|^2}\\ 0 & 0 & -\frac{\sigma_{12}\sigma_{22}^2}{|\boldsymbol{\Sigma}|^2} & -\frac{\sigma_{12}\sigma_{11}^2}{|\boldsymbol{\Sigma}|^2} & \frac{\sigma_{11}^2\sigma_{22}^2+\sigma_{12}^2}{|\boldsymbol{\Sigma}|^2}\end{bmatrix}..$$

3. $c(\boldsymbol{\theta})=\Pr(X_{i1}\le X_{i2})=\Phi\left(\frac{\mu_2-\mu_1}{\sqrt{\sigma_{11}^2\sigma_{22}^2-2\sigma_{12}}}\right).$

4. Let $\dot{c}(\boldsymbol{\theta})=\frac{\partial c(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}$

   Since $\sqrt{n}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})\to_d N(\mathbf{0},I_1(\boldsymbol{\mu},\boldsymbol{\Sigma})^{-1})$ (exactly), the delta method leads to $\sqrt{n}(c(\hat{\boldsymbol{\theta}})-c(\boldsymbol{\theta}))\to_d N(\mathbf{0},\dot{c}(\boldsymbol{\theta})^T I_1(\boldsymbol{\mu},\boldsymbol{\Sigma})^{-1}\dot{c}(\boldsymbol{\theta}))$, and so

   $\frac{\sqrt{n}(c(\hat{\boldsymbol{\theta}})-c(\boldsymbol{\theta}))}{\sqrt{\dot{c}(\boldsymbol{\theta})^T I_1(\boldsymbol{\mu},\boldsymbol{\Sigma})^{-1}\dot{c}(\boldsymbol{\theta})}}$ is the asymptotically pivotal quantity. Hence, the asymptotic confidence interval is

   $$c(\hat{\boldsymbol{\theta}})\pm z_{\alpha/2}\frac{\sqrt{\dot{c}(\boldsymbol{\theta})^T I_1(\boldsymbol{\mu},\boldsymbol{\Sigma})^{-1}\dot{c}(\boldsymbol{\theta})}}{\sqrt{n}}$$