Report#2 Statistical Inference I

Name: Shih Jia-Han

Submitted 2015/10/22; Revised 2015/11/4

Example 5.7 (p.27)

Suppose that (X_i, Y_i), $i = 1, 2, \cdots, n$, is a sample from the bivariate normal density with parameter $\theta = (\xi, \eta, \sigma^2, \tau^2, \rho,)$. The density function is

$$
\begin{aligned}
& f_i(x_i, y_i) \\
&= \frac{1}{2\pi\sigma\tau\sqrt{1-\rho^2}} \exp\left[\frac{-1}{2(1-\rho^2)} \left\{ \left(\frac{x_i - \xi}{\sigma} \right)^2 - 2\rho \left(\frac{x_i - \xi}{\sigma} \right)\left(\frac{y_i - \eta}{\tau} \right) + \left(\frac{y_i - \eta}{\tau} \right)^2 \right\} \right].
\end{aligned}
$$

It follows that

$$
\begin{aligned}
& \frac{1}{2\pi\sigma\tau\sqrt{1-\rho^2}} \exp\left[\frac{-1}{2(1-\rho^2)} \left\{ \left(\frac{x_i - \xi}{\sigma} \right)^2 - 2\rho \left(\frac{x_i - \xi}{\sigma} \right)\left(\frac{y_i - \eta}{\tau} \right) + \left(\frac{y_i - \eta}{\tau} \right)^2 \right\} \right] \\
&= \frac{1}{2\pi} \cdot \frac{1}{\sigma\tau\sqrt{1-\rho^2}} \exp\left[\frac{-1}{2(1-\rho^2)} \left\{ \left(\frac{x_i^2 - 2\xi x_i + \xi^2}{\sigma^2} \right) - \left(\frac{y_i^2 - 2\eta y_i + \eta^2}{\tau^2} \right) \right. \right. \\
&\qquad \left. \left. + 2\rho \left(\frac{x_i y_i - x_i \eta - y_i \xi + \xi\eta}{\sigma\tau} \right) \right\} \right] \\
&= \frac{1}{2\pi} \exp\left[\left(\frac{\xi\tau - \rho\eta\sigma}{(1-\rho^2)\sigma^2\tau} \right) x_i + \left(\frac{-1}{2(1-\rho^2)\sigma^2} \right) x_i^2 \right. \\
&\qquad + \left(\frac{\rho}{(1-\rho^2)\sigma\tau} \right) x_i y_i + \left(\frac{\eta\sigma - \rho\xi\tau}{(1-\rho^2)\sigma\tau^2} \right) y_i + \left(\frac{-1}{2(1-\rho^2)\tau^2} \right) y_i^2 \\
&\qquad \left. - \left\{ \frac{\xi^2\tau^2 + \sigma^2\eta^2 - 2\rho\xi\eta\sigma\tau}{2(1-\rho^2)\sigma^2\tau^2} + \log\sigma + \log\tau + \frac{1}{2}\log(1-\rho) \right\} \right].
\end{aligned}
$$

Therefore,

$$P_\theta(x, y) = \prod_{i=1}^{n} f_i(x_i, y_i)$$

$$= \left(\frac{1}{2\pi}\right)^n \exp\left[\left(\left(\frac{\xi\tau - \rho\eta\sigma}{(1-\rho^2)\sigma^2\tau}\right)\sum_{i=1}^{n} x_i + \left(\frac{-1}{2(1-\rho^2)\sigma^2}\right)\sum_{i=1}^{n} x_i^2\right.\right.$$

$$+\left.\left(\left(\frac{\rho}{(1-\rho^2)\sigma\tau}\right)\right)\sum_{i=1}^{n} x_i y_i + \left(\frac{\eta\sigma - \rho\xi\tau}{(1-\rho^2)\sigma\tau^2}\right)\sum_{i=1}^{n} y_i$$

$$+\left(\frac{-1}{2(1-\rho^2)\tau^2}\right)\sum_{i=1}^{n} y_i^2$$

$$\left. - n\left\{\frac{\xi^2\tau^2 + \sigma^2\eta^2 - 2\rho\xi\eta\sigma\tau}{2(1-\rho^2)\sigma^2\tau^2} + \log\sigma + \log\tau + \frac{1}{2}\log(1-\rho)\right\}\right].$$

Hence

$$P_\theta(x, y) = \exp\left\{\sum_{i=1}^{5} \eta_i(\theta) T_i(x, y) - A(\eta)\right\} h(x)$$

is the canonical form of a five-dimensional exponential family, where

$$\eta_1(\theta) = \frac{\xi\tau - \rho\eta\sigma}{(1-\rho^2)\sigma^2\tau},\ \ \eta_2(\theta) = \frac{-1}{2(1-\rho^2)\sigma^2},\ \ \eta_3(\theta) = \frac{\rho}{(1-\rho^2)\sigma\tau},$$

$$\eta_4(\theta) = \frac{\eta\sigma - \rho\xi\tau}{(1-\rho^2)\sigma\tau^2},\ \ \eta_5(\theta) = \frac{-1}{2(1-\rho^2)\tau^2},\ \ T_1(x, y) = \sum_{i=1}^{n} x_i,$$

$$T_2(x, y) = \sum_{i=1}^{n} x_i^2,\ \ T_3(x, y) = \sum_{i=1}^{n} x_i y_i,\ \ T_4(x, y) = \sum_{i=1}^{n} y_i,$$

$$T_5(x, y) = \sum_{i=1}^{n} y_i^2,\ \ h(x, y) = \left(\frac{1}{2\pi}\right)^n,$$

$$A(\eta) = n\left\{\frac{\xi^2\tau^2 + \sigma^2\eta^2 - 2\rho\xi\eta\sigma\tau}{2(1-\rho^2)\sigma^2\tau^2} + \log\sigma + \log\tau + \frac{1}{2}\log(1-\rho)\right\},$$

$$\chi = (-\infty, \infty)^2 \times \cdots \times (-\infty, \infty)^2 \ \text{ and}$$

$$\Theta = \{(\eta_1, \eta_2, \eta_3, \eta_4, \eta_5); \eta_1 \in R, \eta_2 < 0, \eta_3 \in R, \eta_4 < 0, \eta_5 \in R\}.$$

Problem 5.3 (p.66)

Show that the distribution of a sample from p-variate normal density (4.15) constitutes an s-dimensional exponential family. Determine s and identify the functions η_i, T_i and B of (5.1).

Solution:

$$f(x_1,\cdots,x_p)=\frac{|\mathbf{B}|^{-1}}{\sqrt{(2\pi)^p}}\exp\left\{\frac{-1}{2}(\mathbf{x}-\mathbf{a})'\Sigma^{-1}(\mathbf{x}-\mathbf{a})\right\}.$$

Since

$$\mathbf{\Sigma}=\begin{bmatrix}\sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1p}\\ \sigma_{21} & \sigma_2^2 & & \sigma_{2p}\\ \vdots & & \ddots & \vdots\\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_p^2\end{bmatrix}.$$

Let $\mathbf{\Sigma}^{-1}=\mathbf{\Lambda}$, we have $\mathbf{\Lambda}=\mathbf{\Lambda}'$.

Therefore,

$$\begin{aligned}
f(x_1,\cdots,x_p)&=\frac{|\mathbf{B}|^{-1}}{\sqrt{(2\pi)^p}}\exp\left\{\frac{-1}{2}(\mathbf{x}-\mathbf{a})'\mathbf{\Sigma}^{-1}(\mathbf{x}-\mathbf{a})\right\}\\
&=\frac{|\mathbf{B}|^{-1}}{\sqrt{(2\pi)^p}}\exp\left\{\frac{-1}{2}(\mathbf{x}-\mathbf{a})'\mathbf{\Lambda}(\mathbf{x}-\mathbf{a})\right\}\\
&=\left(\frac{1}{2\pi}\right)^{\frac{p}{2}}|\mathbf{B}|^{-1}\exp\left\{\frac{-1}{2}(\mathbf{x}'\mathbf{\Lambda}\mathbf{x}-\mathbf{x}'\mathbf{\Lambda}\mathbf{a}-\mathbf{a}'\mathbf{\Lambda}\mathbf{x}+\mathbf{a}'\mathbf{\Lambda}\mathbf{a})\right\}\\
&=\left(\frac{1}{2\pi}\right)^{\frac{p}{2}}|\mathbf{B}|^{-1}\exp\left\{\mathbf{a}'\mathbf{\Lambda}\mathbf{x}-\frac{1}{2}\mathbf{x}'\mathbf{\Lambda}\mathbf{x}-\mathbf{a}'\mathbf{\Lambda}\mathbf{a}\right\}\\
&=\left(\frac{1}{2\pi}\right)^{\frac{p}{2}}\exp\left\{\mathbf{a}'\mathbf{\Lambda}\mathbf{x}-\frac{1}{2}\mathbf{x}'\mathbf{\Lambda}\mathbf{x}-(\mathbf{a}'\mathbf{\Lambda}\mathbf{a}+\log|\mathbf{B}|)\right\}\\
&=\left(\frac{1}{2\pi}\right)^{\frac{p}{2}}\exp\left\{\mathbf{a}'\mathbf{\Sigma}^{-1}\mathbf{x}-\frac{1}{2}\mathbf{x}'\mathbf{\Sigma}^{-1}\mathbf{x}-(\mathbf{a}'\mathbf{\Sigma}^{-1}\mathbf{a}+\log|\mathbf{B}|)\right\}.
\end{aligned}$$

Let

$$\boldsymbol{\Sigma}^{-1} = \begin{bmatrix} \boldsymbol{\Sigma}_1^{-1} & \boldsymbol{\Sigma}_2^{-1} & \cdots & \boldsymbol{\Sigma}_p^{-1} \end{bmatrix}.$$

Hence

$$\mathbf{a}'\boldsymbol{\Sigma}^{-1}\mathbf{x} = \mathbf{a}'\begin{bmatrix} \boldsymbol{\Sigma}_1^{-1} & \boldsymbol{\Sigma}_2^{-1} & \cdots & \boldsymbol{\Sigma}_p^{-1} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} = \begin{bmatrix} \mathbf{a}'\boldsymbol{\Sigma}_1^{-1} & \mathbf{a}'\boldsymbol{\Sigma}_2^{-1} & \cdots & \mathbf{a}'\boldsymbol{\Sigma}_p^{-1} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}$$

$$= \mathbf{a}'\boldsymbol{\Sigma}_1^{-1}x_1 + \mathbf{a}'\boldsymbol{\Sigma}_2^{-1}x_2 + \cdots \mathbf{a}'\boldsymbol{\Sigma}_p^{-1}x_p$$

$$= \sum_{i=1}^{p} \mathbf{a}'\boldsymbol{\Sigma}_i^{-1}x_i.$$

Let

$$\boldsymbol{\Sigma}^{-1} = \begin{bmatrix} \boldsymbol{\Sigma}_{11}^{-1} & \boldsymbol{\Sigma}_{12}^{-1} & \cdots & \boldsymbol{\Sigma}_{1p}^{-1} \\ \boldsymbol{\Sigma}_{12}^{-1} & \boldsymbol{\Sigma}_{22}^{-1} & & \boldsymbol{\Sigma}_{2p}^{-1} \\ \vdots & & \ddots & \vdots \\ \boldsymbol{\Sigma}_{1p}^{-1} & \boldsymbol{\Sigma}_{2p}^{-1} & \cdots & \boldsymbol{\Sigma}_{pp}^{-1} \end{bmatrix}.$$

Hence

$$\mathbf{x}'\boldsymbol{\Sigma}^{-1}\mathbf{x} = \boldsymbol{\Sigma}_{11}^{-1}x_1^2 + \boldsymbol{\Sigma}_{22}^{-1}x_2^2 + \cdots + \boldsymbol{\Sigma}_{pp}^{-1}x_p^2$$

$$+ 2\boldsymbol{\Sigma}_{12}^{-1}x_1x_2 + 2\boldsymbol{\Sigma}_{13}^{-1}x_1x_3 + \cdots + 2\boldsymbol{\Sigma}_{(p-1)p}^{-1}x_{p-1}x_p$$

$$= \sum_{i=1}^{p} \boldsymbol{\Sigma}_{ii}^{-1}x_i^2 + \sum_{i<j} 2\boldsymbol{\Sigma}_{ij}^{-1}x_ix_j.$$

Therefore,

$$P(x) = \left(\frac{1}{2\pi}\right)^{\frac{p}{2}} \exp\left\{\mathbf{a}'\boldsymbol{\Sigma}^{-1}\mathbf{x} - \frac{1}{2}\mathbf{x}'\boldsymbol{\Sigma}^{-1}\mathbf{x} - (\mathbf{a}'\boldsymbol{\Sigma}^{-1}\mathbf{a} + \log|\mathbf{B}|)\right\}$$

$$= \left(\frac{1}{2\pi}\right)^{\frac{p}{2}} \exp\left\{\sum_{i=1}^{p} \mathbf{a}'\boldsymbol{\Sigma}_i^{-1}x_i - \frac{1}{2}\sum_{i=1}^{p} \boldsymbol{\Sigma}_{ii}^{-1}x_i^2 - \sum_{i<j} \boldsymbol{\Sigma}_{ij}^{-1}x_ix_j - (\mathbf{a}'\boldsymbol{\Sigma}^{-1}\mathbf{a} + \log|\mathbf{B}|)\right\},$$

is a $(p^2 + 3p)/2$-dimensional exponential family.

Therefore, it can be written as

$$P(x) = \exp\{\eta_i T_i - B\} h(x),$$

where

$$T_1 = x_1, T_2 = x_2, \cdots, T_p = x_p, T_{p+1} = x_1^2, T_{p+2} = x_2^2, \cdots, T_{2p} = x_p^2,$$

$$T_{2p+1} = x_1 x_2, T_{2p+2} = x_1 x_3, \cdots, T_{(p^2+3p)/2} = x_{p-1} x_p,$$

$$\eta_1 = \mathbf{a}'\mathbf{\Sigma}_1^{-1}, \eta_2 = \mathbf{a}'\mathbf{\Sigma}_2^{-1}, \cdots, \eta_p = \mathbf{a}'\mathbf{\Sigma}_p^{-1}, \eta_{p+1} = \frac{-1}{2}\mathbf{\Sigma}_{11}^{-1}, \eta_{p+2} = \frac{-1}{2}\mathbf{\Sigma}_{22}^{-1}, \cdots, \eta_{2p} = \frac{-1}{2}\mathbf{\Sigma}_{pp}^{-1}$$

$$\eta_{2p+1} = -\mathbf{\Sigma}_{12}^{-1}, \eta_{2p+2} = -\mathbf{\Sigma}_{13}^{-1}, \cdots, \eta_{(p^2+3p)/2} = -\mathbf{\Sigma}_{(p-1)p}^{-1}$$

$$B = \mathbf{a}'\mathbf{\Sigma}^{-1}\mathbf{a} + \log|\mathbf{B}|.$$

Problem 5.6 (p.66)

In the density (5.1)

(a) For $s=1$, show that

$$E_\theta[T(X)] = B'(\theta)/\eta'(\theta) \quad \text{and} \quad \text{var}_\theta[T(X)] = \frac{B''(\theta\quad)}{\eta'(\theta)^2} - \frac{\eta''(\theta)B'(\theta)}{\eta'(\theta)^3}.$$

(b) For $s>1$, show that

$$E_\theta[T(X)] = J^{-1}\nabla B,$$

where J is the Jacobian matrix defined by $J = \partial\eta_j / \partial\theta_i$ and ∇B is the gradient vector $\nabla B = \partial B(\theta)/\partial\theta_i$.

Solution of (a):

By density (5.1) and $s=1$,

$$\int_\chi \exp\{\eta_1(\theta)T_1(x) - B(\theta)\}h(x)d\mu(x) = 1$$

$$\Rightarrow \frac{\partial}{\partial\theta}\int_\chi \exp\{\eta_1(\theta)T_1(x) - B(\theta)\}h(x)d\mu(x) = 0$$

$$\Rightarrow \int_\chi \frac{\partial}{\partial\theta}\exp\{\eta_1(\theta)T_1(x) - B(\theta)\}h(x)d\mu(x) = 0$$

$$\Rightarrow \int_\chi \{\eta_1'(\theta)T_1(x) - B'(\theta)\}\exp\{\eta_1(\theta)T_1(x) - B(\theta)\}h(x)d\mu(x) = 0$$

$$\Rightarrow \eta_1'(\theta)\int_\chi T_1(x)\exp\{\eta_1(\theta)T_1(x) - B(\theta)\}h(x)d\mu(x) = B'(\theta)$$

$$\Rightarrow \eta_1'(\theta)E_\theta[T_1(x)] = B'(\theta)$$

$$\Rightarrow E_\theta[T_1(x)] = \frac{B'(\theta)}{\eta_1'(\theta)}.$$

Similarly,

$$\int_{\chi} \exp\{\eta_1(\theta)T_1(x) - B(\theta)\}h(x)d\mu(x) = 1$$

$$\Rightarrow \frac{\partial^2}{\partial\theta^2}\int_{\chi} \exp\{\eta_1(\theta)T_1(x) - B(\theta)\}h(x)d\mu(x) = 0$$

$$\Rightarrow \int_{\chi} \frac{\partial^2}{\partial\theta^2} \exp\{\eta_1(\theta)T_1(x) - B(\theta)\}h(x)d\mu(x) = 0$$

$$\Rightarrow \int_{\chi} \frac{\partial}{\partial\theta}[\{\eta_1'(\theta)T_1(x) - B'(\theta)\}\exp\{\eta_1(\theta)T_1(x) - B(\theta)\}]h(x)d\mu(x) = 0$$

$$\Rightarrow \int_{\chi} \frac{\partial}{\partial\theta}[\eta_1'(\theta)T_1(x)\exp\{\eta_1(\theta)T_1(x) - B(\theta)\}]h(x)d\mu(x) = B''(\theta)$$

$$\Rightarrow \int_{\chi} [\{\eta_1''(\theta)T_1(x) + \eta_1'(\theta)^2 T_1(x) - \eta_1'(\theta)B'(\theta)\}$$

$$\times \exp\{\eta_1(\theta)T_1(x) - B(\theta)\}]h(x)d\mu(x) = B''(\theta)$$

$$\Rightarrow \int_{\chi} \eta_1''(\theta)T_1(x)\exp\{\eta_1(\theta)T_1(x) - B(\theta)\}]h(x)d\mu(x)$$

$$+\int_{\chi} \{\eta_1'(\theta)^2 T_1(x) - \eta_1'(\theta)B'(\theta)\}\exp\{\eta_1(\theta)T_1(x) - B(\theta)\}]h(x)d\mu(x) = B''(\theta)$$

$$\Rightarrow \frac{\eta_1''(\theta)}{\eta_1'(\theta)^2} E_\theta[T_1(x)]$$

$$+\int_{\chi} \left\{T_1(x) - \frac{B'(\theta)}{\eta_1'(\theta)}\right\}\exp\{\eta_1(\theta)T_1(x) - B(\theta)\}]h(x)d\mu(x) = \frac{B''(\theta)}{\eta_1'(\theta)^2}$$

$$\Rightarrow \frac{B'(\theta)\eta_1''(\theta)}{\eta_1'(\theta)^3}$$

$$+\int_{\chi} \{T_1(x) - E_\theta[T_1(x)]\}\exp\{\eta_1(\theta)T_1(x) - B(\theta)\}]h(x)d\mu(x) = \frac{B''(\theta)}{\eta_1'(\theta)^2}$$

$$\Rightarrow \frac{B'(\theta)\eta_1''(\theta)}{\eta_1'(\theta)^3} + \mathrm{var}_\theta[T_1(x)] = \frac{B''(\theta)}{\eta_1'(\theta)^2}$$

$$\Rightarrow \mathrm{var}_\theta[T_1(x)] = \frac{B''(\theta)}{\eta_1'(\theta)^2} - \frac{B'(\theta)\eta_1''(\theta)}{\eta_1'(\theta)^3}.$$

Solution of (b):

By density (5.1) and $s > 1$,

$$\int_{\chi} \exp\left\{\sum_{i=1}^{s} \eta_i(\theta) T_i(x) - B(\theta)\right\} h(x) d\mu(x) = 1$$

$$\Rightarrow \frac{\partial}{\partial \theta_i} \int_{\chi} \exp\left\{\sum_{i=1}^{s} \eta_i(\theta) T_i(x) - B(\theta)\right\} h(x) d\mu(x) = 0$$

$$\Rightarrow \int_{\chi} \frac{\partial}{\partial \theta_i} \exp\left\{\sum_{i=1}^{s} \eta_i(\theta) T_i(x) - B(\theta)\right\} h(x) d\mu(x) = 0$$

$$\Rightarrow \int_{\chi} \left\{\frac{\partial}{\partial \theta_i} \sum_{i=1}^{s} \eta_i(\theta) T_i(x) - \frac{\partial B(\theta)}{\partial \theta_i}\right\} \exp\left\{\sum_{i=1}^{s} \eta_i(\theta) T_i(x) - B(\theta)\right\} h(x) d\mu(x) = 0$$

$$\Rightarrow \int_{\chi} \left\{\frac{\partial \eta_1(\theta)}{\partial \theta_i} T_1(x) + \cdots + \frac{\partial \eta_s(\theta)}{\partial \theta_i} T_s(x)\right\} \exp\left\{\sum_{i=1}^{s} \eta_i(\theta) T_i(x) - B(\theta)\right\} h(x) d\mu(x) = \frac{\partial B(\theta)}{\partial \theta_i}$$

$$\Rightarrow \begin{bmatrix} \frac{\partial \eta_1}{\partial \theta_i} & \frac{\partial \eta_2}{\partial \theta_i} & \cdots & \frac{\partial \eta_s}{\partial \theta_i} \end{bmatrix} \begin{bmatrix} E_\theta[T_1(x)] \\ E_\theta[T_2(x)] \\ \vdots \\ E_\theta[T_s(x)] \end{bmatrix} = \frac{\partial B(\theta)}{\partial \theta_i}.$$

Therefore,

$$\begin{bmatrix} \frac{\partial \eta_1}{\partial \theta_1} & \frac{\partial \eta_2}{\partial \theta_1} & \cdots & \frac{\partial \eta_s}{\partial \theta_1} \\ \frac{\partial \eta_1}{\partial \theta_2} & \frac{\partial \eta_s}{\partial \theta_2} & \cdots & \frac{\partial \eta_s}{\partial \theta_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \eta_1}{\partial \theta_s} & \frac{\partial \eta_2}{\partial \theta_s} & \cdots & \frac{\partial \eta_s}{\partial \theta_s} \end{bmatrix} \begin{bmatrix} E_\theta[T_1(x)] \\ E_\theta[T_2(x)] \\ \vdots \\ E_\theta[T_s(x)] \end{bmatrix} = \begin{bmatrix} \frac{\partial B(\theta)}{\partial \theta_1} \\ \frac{\partial B(\theta)}{\partial \theta_2} \\ \vdots \\ \frac{\partial B(\theta)}{\partial \theta_s} \end{bmatrix}$$

$$\Rightarrow J \times E_\theta[T(x)] = \nabla B$$

$$\Rightarrow E_\theta[T(x)] = J^{-1} \nabla B.$$