556: MATHEMATICAL STATISTICS I

DISTRIBUTIONS DERIVED FROM NORMAL RANDOM SAMPLES

Suppose that $X_1, \ldots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ form a random sample.

(a) By the univariate transformation result

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$$

(b) By the multivariate transformation result

$$T = \frac{\overline{X} - \mu}{s/\sqrt{n}} \sim \text{Student}(n-1)$$

as T can be written $T = Z/\sqrt{V/\nu}$, where $Z \sim \mathcal{N}(0,1)$ and $V \sim \chi^2_\nu$ are independent random variables defined by

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \qquad V = \frac{(n-1)s^2}{\sigma^2}$$

(c) **Fisher-F distribution**: By the multivariate transformation result, if $V_1 \sim \chi^2_{k_1}$ and $V_2 \sim \chi^2_{k_2}$ are independent random variables, then

$$Q = \frac{V_1/k_1}{V_2/k_2}$$

has pdf

$$f_Q(x) = \frac{\Gamma\left(\dfrac{k_1+k_2}{2}\right)}{\Gamma\left(\dfrac{k_1}{2}\right)\Gamma\left(\dfrac{k_2}{2}\right)} \left(\frac{k_1}{k_2}\right)^{k_1/2} x^{k_1/2-1} \left(1 + \frac{k_1}{k_2}x\right)^{-(k_1+k_2)/2} \qquad x > 0$$

and zero otherwise. We say that $Q \sim \text{Fisher}(k_1, k_2)$.

Now, if $X_1, \ldots, X_{n_X} \sim \mathcal{N}(\mu_X, \sigma_X^2)$ and $Y_1, \ldots, Y_{n_Y} \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ are independent random samples from different distributions, then as

$$\frac{(n_X-1)s_X^2}{\sigma_X^2} \sim \chi^2_{n_X-1} \qquad \frac{(n_Y-1)s_Y^2}{\sigma_Y^2} \sim \chi^2_{n_Y-1}$$

it follows that

$$\frac{s_X^2/\sigma_X^2}{s_Y^2/\sigma_Y^2} \sim \text{Fisher}(n_X-1, n_Y-1)$$

Note that if $X \sim \text{Fisher}(k_1, k_2)$, then

$$\mathbb{E}_X[X] = \frac{k_2}{k_2-2} \ \ (\text{if } k_2 > 2) \qquad \text{Var}_X[X] = 2\left(\frac{k_2}{k_2-2}\right)^2 \frac{(k_1+k_2-2)}{k_1(k_2-4)} \ \ (\text{if } k_2 > 4)$$

Also

$$X \sim \text{Fisher}(k_1, k_2) \implies \frac{1}{X} \sim \text{Fisher}(k_2, k_1)$$

$$X \sim \text{Student}(k_1) \implies X^2 \sim \text{Fisher}(1, k_1)$$

$$X \sim \text{Fisher}(k_1, k_2) \implies \frac{(k_1/k_2)X}{1+(k_1/k_2)X} \sim \text{Beta}\left(\frac{k_1}{2}, \frac{k_2}{2}\right)$$

Suppose that vector random variable $\mathbf{X} = (X_1, X_2, \dots, X_k)^\top$ has a multivariate normal distribution with pdf given by

$$f_{\mathbf{X}}(\mathbf{x}) = \left(\frac{1}{2\pi}\right)^{k/2} \frac{1}{|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}\mathbf{x}^\top \Sigma^{-1}\mathbf{x}\right\} \tag{1}$$

where Σ is the $k \times k$ variance-covariance matrix (we can consider here the case where the expected value $\underset{\sim}{\mu}$ is the $k \times 1$ zero vector; results for the general case are easily available by transformation).

Consider partitioning $\mathbf{X}$ into two components $\mathbf{X}_1$ and $\mathbf{X}_2$ of dimensions k_1 and $k_2 = k - k_1$ respectively, that is,

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix}.$$

We attempt to deduce

(a) the **marginal** distribution of $\mathbf{X}_1$, and
(b) the **conditional** distribution of $\mathbf{X}_2$ **given** that $\mathbf{X}_1 = \mathbf{x}_1$.

First, write

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

where Σ_{11} is $k_1 \times k_1$, Σ_{22} is $k_2 \times k_2$, $\Sigma_{21} = \Sigma_{12}^\top$, and

$$\Sigma^{-1} = \mathbf{V} = \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{bmatrix}$$

so that $\Sigma\mathbf{V} = \mathbf{I}_k$ ($\mathbf{I}_r$ is the $r \times r$ identity matrix) gives

$$\begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{k_1} & 0 \\ 0 & \mathbf{I}_{k_2} \end{bmatrix}$$

where 0 represents the zero matrix of appropriate dimension. More specifically,

$$\Sigma_{11}\mathbf{V}_{11} + \Sigma_{12}\mathbf{V}_{21} = \mathbf{I}_{k_1} \tag{2}$$
$$\Sigma_{11}\mathbf{V}_{12} + \Sigma_{12}\mathbf{V}_{22} = 0 \tag{3}$$
$$\Sigma_{21}\mathbf{V}_{11} + \Sigma_{22}\mathbf{V}_{21} = 0 \tag{4}$$
$$\Sigma_{21}\mathbf{V}_{12} + \Sigma_{22}\mathbf{V}_{22} = \mathbf{I}_{k_2}. \tag{5}$$

From the multivariate normal pdf in equation (1), we can re-express the term in the exponent as

$$\mathbf{x}^\top \Sigma^{-1}\mathbf{x} = \mathbf{x}_1^\top \mathbf{V}_{11}\mathbf{x}_1 + \mathbf{x}_1^\top \mathbf{V}_{12}\mathbf{x}_2 + \mathbf{x}_2^\top \mathbf{V}_{21}\mathbf{x}_1 + \mathbf{x}_2^\top \mathbf{V}_{22}\mathbf{x}_2. \tag{6}$$

In order to compute the marginal and conditional distributions, we must complete the square in $\mathbf{x}_2$ in this expression. We can write

$$\mathbf{x}^\top \Sigma^{-1}\mathbf{x} = (\mathbf{x}_2 - \mathbf{m})^\top \mathbf{M}(\mathbf{x}_2 - \mathbf{m}) + \mathbf{c} \tag{7}$$

and by comparing with equation (6) we can deduce that, for quadratic terms in $\mathbf{x}_2$,

$$\mathbf{x}_2^\top \mathbf{V}_{22}\mathbf{x}_2 = \mathbf{x}_2^\top \mathbf{M}\mathbf{x}_2 \qquad \therefore \qquad \mathbf{M} = \mathbf{V}_{22} \tag{8}$$

for linear terms

$$\mathbf{x}_2^\top \mathbf{V}_{21}\mathbf{x}_1 = -\mathbf{x}_2^\top \mathbf{M}\mathbf{m} \qquad \therefore \qquad \mathbf{m} = -\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{x}_1 \tag{9}$$

and for constant terms

$$\mathbf{x}_1^\top \mathbf{V}_{11}\mathbf{x}_1 = \mathbf{c} + \mathbf{m}^\top \mathbf{M}\mathbf{m} \qquad \therefore \qquad \mathbf{c} = \mathbf{x}_1^\top(\mathbf{V}_{11} - \mathbf{V}_{21}^\top \mathbf{V}_{22}^{-1}\mathbf{V}_{21})\mathbf{x}_1 \tag{10}$$

thus yielding all the terms required for equation (7), that is

$$\mathbf{x}^\top \Sigma^{-1}\mathbf{x} = (\mathbf{x}_2 + \mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{x}_1)^\top \mathbf{V}_{22}(\mathbf{x}_2 + \mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{x}_1) + \mathbf{x}_1^\top(\mathbf{V}_{11} - \mathbf{V}_{21}^\top \mathbf{V}_{22}^{-1}\mathbf{V}_{21})\mathbf{x}_1, \tag{11}$$

which, crucially, is a sum of two terms, where the first can be interpreted as a function of $\mathbf{x}_2$, given $\mathbf{x}_1$, and the second is a function of $\mathbf{x}_1$ only.

Hence we have an immediate factorization of the full joint pdf using the chain rule for random variables;

$$f_{\mathbf{X}}(\mathbf{x}) = f_{\mathbf{X}_2|\mathbf{X}_1}(\mathbf{x}_2|\mathbf{x}_1) f_{\mathbf{X}_1}(\mathbf{x}_1) \tag{12}$$

where

$$f_{\mathbf{X}_2|\mathbf{X}_1}(\mathbf{x}_2|\mathbf{x}_1) \propto \exp\left\{-\frac{1}{2}(\mathbf{x}_2 + \mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{x}_1)^\top \mathbf{V}_{22}(\mathbf{x}_2 + \mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{x}_1)\right\} \tag{13}$$

giving that

$$\mathbf{X}_2|\mathbf{X}_1 = \mathbf{x}_1 \ \sim \mathcal{N}_{k_2}\left(-\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{x}_1, \mathbf{V}_{22}^{-1}\right) \tag{14}$$

and

$$f_{\mathbf{X}_1}(\mathbf{x}_1) \propto \exp\left\{-\frac{1}{2}\mathbf{x}_1^\top(\mathbf{V}_{11} - \mathbf{V}_{21}^\top \mathbf{V}_{22}^{-1}\mathbf{V}_{21})\mathbf{x}_1\right\} \tag{15}$$

giving that

$$\mathbf{X}_1 \sim \mathcal{N}_{k_1}\left(0, (\mathbf{V}_{11} - \mathbf{V}_{21}^\top \mathbf{V}_{22}^{-1}\mathbf{V}_{21})^{-1}\right). \tag{16}$$

But, from equation (3), $\Sigma_{12} = -\Sigma_{11}\mathbf{V}_{12}\mathbf{V}_{22}^{-1}$, and then from equation (2), substituting in Σ_{12},

$$\Sigma_{11}\mathbf{V}_{11} - \Sigma_{11}\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21} = \mathbf{I}_{k_1} \qquad \therefore \qquad \Sigma_{11} = (\mathbf{V}_{11} - \mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21})^{-1} = (\mathbf{V}_{11} - \mathbf{V}_{21}^\top \mathbf{V}_{22}^{-1}\mathbf{V}_{21})^{-1}.$$

Hence, by inspection of equation (16), we conclude that

$$\boxed{\mathbf{X}_1 \sim \mathcal{N}_{k_1}(0, \Sigma_{11}),} \tag{17}$$

that is, we can extract the Σ_{11} block of Σ to define the marginal sigma matrix of $\mathbf{X}_1$.

Using similar arguments, we can define the conditional distribution from equation (14) more precisely. First, from equation (3), $\mathbf{V}_{12} = -\Sigma_{11}^{-1}\Sigma_{12}\mathbf{V}_{22}$, and then from equation (5), substituting in $\mathbf{V}_{12}$

$$-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\mathbf{V}_{22} + \Sigma_{22}\mathbf{V}_{22} = \mathbf{I}_{k-k_1} \qquad \therefore \qquad \mathbf{V}_{22}^{-1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} = \Sigma_{22} - \Sigma_{12}^\top \Sigma_{11}^{-1}\Sigma_{12}.$$

Finally, from equation (3), taking transposes on both sides, we have that $\mathbf{V}_{21}\Sigma_{11} + \mathbf{V}_{22}\Sigma_{21} = 0$. Then pre-multiplying by $\mathbf{V}_{22}^{-1}$, and post-multiplying by Σ_{11}^{-1}, we have

$$\mathbf{V}_{22}^{-1}\mathbf{V}_{21} + \Sigma_{21}\Sigma_{11}^{-1} = 0 \qquad \therefore \qquad \mathbf{V}_{22}^{-1}\mathbf{V}_{21} = -\Sigma_{21}\Sigma_{11}^{-1},$$

so we have, substituting into equation (14), that

$$\boxed{\mathbf{X}_2|\mathbf{X}_1 = \mathbf{x}_1 \ \sim \mathcal{N}_{k_2}\left(\Sigma_{21}\Sigma_{11}^{-1}\mathbf{x}_1, \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right).} \tag{18}$$

Thus any marginal, and any conditional distribution of a multivariate normal joint distribution is also multivariate normal, as the choices of $\mathbf{X}_1$ and $\mathbf{X}_2$ are arbitrary.