NORMAL DISTRIBUTION FACTSHEET

The details of the PROOFS of these results are not examinable in M2S1. However, you should know the results themselves.

MULTIVARIATE NORMAL DISTRIBUTION: MARGINALS AND CONDITIONALS

Suppose that vector random variable $\underset{\sim}{X} = (X_1, X_2, \ldots, X_k)^\mathsf{T}$ has a multivariate normal distribution with pdf given by

$$f_{\underset{\sim}{X}}(\underset{\sim}{x}) = \left(\frac{1}{2\pi}\right)^{k/2} \frac{1}{|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}\underset{\sim}{x}^\mathsf{T}\Sigma^{-1}\underset{\sim}{x}\right\} \qquad (1)$$

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 $\underset{\sim}{X}$ into two components $\underset{\sim}{X}_1$ and $\underset{\sim}{X}_2$ of dimensions d and $k-d$ respectively, that is,

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

We attempt to deduce

(a) the marginal distribution of $\underset{\sim}{X}_1$, and

(b) the conditional distribution of $\underset{\sim}{X}_2$ given that $\underset{\sim}{X}_1 = \underset{\sim}{x}_1$.

First, write

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

where Σ_{11} is $d \times d$, Σ_{22} is $(k-d) \times (k-d)$, $\Sigma_{21} = \Sigma_{12}^\mathsf{T}$, and

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

so that $\Sigma V = I_k$ (I_r is the $r \times r$ identity matrix) gives

$$\begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} = \begin{bmatrix} I_d & 0 \\ 0 & I_{k-d} \end{bmatrix}$$

and more specifically the four relations

$$\begin{aligned} \Sigma_{11}V_{11} + \Sigma_{12}V_{21} &= I_d && (2) \\ \Sigma_{11}V_{12} + \Sigma_{12}V_{22} &= 0 && (3) \\ \Sigma_{21}V_{11} + \Sigma_{22}V_{21} &= 0 && (4) \\ \Sigma_{21}V_{12} + \Sigma_{22}V_{22} &= I_{k-d}. && (5) \end{aligned}$$

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

$$\underset{\sim}{x}^\mathsf{T}\Sigma^{-1}\underset{\sim}{x} = \underset{\sim}{x}_1^\mathsf{T}V_{11}\underset{\sim}{x}_1 + \underset{\sim}{x}_1^\mathsf{T}V_{12}\underset{\sim}{x}_2 + \underset{\sim}{x}_2^\mathsf{T}V_{21}\underset{\sim}{x}_1 + \underset{\sim}{x}_2^\mathsf{T}V_{22}\underset{\sim}{x}_2. \qquad (6)$$

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

$$\underset{\sim}{x}^\mathsf{T}\Sigma^{-1}\underset{\sim}{x} = (\underset{\sim}{x}_2 - \underset{\sim}{m})^\mathsf{T}M(\underset{\sim}{x}_2 - \underset{\sim}{m}) + \underset{\sim}{c} \qquad (7)$$

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

$$\underset{\sim}{x}_2^\top V_{22} \underset{\sim}{x}_2 = \underset{\sim}{x}_2^\top M \underset{\sim}{x}_2 \qquad \therefore \qquad M = V_{22} \tag{8}$$

for linear terms

$$\underset{\sim}{x}_2^\top V_{21} \underset{\sim}{x}_1 = -\underset{\sim}{x}_2^\top M \underset{\sim}{m} \qquad \therefore \qquad \underset{\sim}{m} = -V_{22}^{-1} V_{21} \underset{\sim}{x}_1 \tag{9}$$

and for constant terms

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

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

$$\underset{\sim}{x}^\top \Sigma^{-1} \underset{\sim}{x} = (\underset{\sim}{x}_2 + V_{22}^{-1} V_{21} \underset{\sim}{x}_1)^\top V_{22} (\underset{\sim}{x}_2 + V_{22}^{-1} V_{21} \underset{\sim}{x}_1) + \underset{\sim}{x}_1^\top (V_{11} - V_{21}^\top V_{22}^{-1} V_{21}) \underset{\sim}{x}_1, \tag{11}$$

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

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

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

where

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

giving that

$$\underset{\sim}{X}_2|\underset{\sim}{X}_1 = \underset{\sim}{x}_1 \sim N\left(-V_{22}^{-1} V_{21} \underset{\sim}{x}_1, V_{22}^{-1} \right) \tag{14}$$

and

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

giving that

$$\underset{\sim}{X}_1 \sim N\left(0, (V_{11} - V_{21}^\top V_{22}^{-1} V_{21})^{-1} \right). \tag{16}$$

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

$$\Sigma_{11} V_{11} - \Sigma_{11} V_{12} V_{22}^{-1} V_{21} = I_d \qquad \therefore \qquad \Sigma_{11} = (V_{11} - V_{12} V_{22}^{-1} V_{21})^{-1} = (V_{11} - V_{21}^\top V_{22}^{-1} V_{21})^{-1}.$$

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

$$\boxed{\underset{\sim}{X}_1 \sim N(0, \Sigma_{11}),} \tag{17}$$

that is, we can extract the Σ_{11} block of Σ to define the marginal variance-covariance matrix of $\underset{\sim}{X}_1$.

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

$$-\Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} V_{22} + \Sigma_{22} V_{22} = I_{k-d} \qquad \therefore \qquad V_{22}^{-1} = \Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} = \Sigma_{22} - \Sigma_{12}^\top \Sigma_{11}^{-1} \Sigma_{12}.$$

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

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

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

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

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

SAMPLING DISTRIBUTION FOR NORMAL SAMPLES

THEOREM If $X_1, ..., X_n$ is a random sample from a normal distribution, say $X_i \sim N(\mu, \sigma^2)$, then
(a) $\overline{X}$ is independent of $\left\{X_i - \overline{X}, i = 1, ..., n\right\}$
(b) $\overline{X}$ and s^2 are independent random variables
(c) The random variable

$$\frac{(n-1)s^2}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$$

has a **chi-squared distribution** with $n-1$ degrees of freedom.

Proof: (a) The joint pdf $X_1, ..., X_n$ is the multivariate normal density

$$f_{X_1,...,X_n}(x_1, ..., x_n) = \left(\frac{1}{2\pi}\right)^{n/2}\frac{1}{|\Sigma|^{1/2}}\exp\left\{-\frac{1}{2}(\underset{\sim}{x} - \underset{\sim}{\mu})^T\Sigma^{-1}(\underset{\sim}{x} - \underset{\sim}{\mu})\right\}$$

where $\Sigma = \sigma^2 I_n$, and I_n is the $n \times n$ identity matrix. Consider the multivariate transformation to $Y_1, ..., Y_n$ where

$$\left.\begin{aligned} Y_1 &= \overline{X} \\ Y_i &= X_i - \overline{X},\ i = 2, ..., n \end{aligned}\right\} \Longleftrightarrow \left\{\begin{aligned} X_1 &= Y_1 - \sum_{i=2}^{n} Y_i \\ X_i &= Y_i + Y_1, \quad i = 2, ..., n \end{aligned}\right.$$

Thus, in vector terms $\underset{\sim}{Y} = A\underset{\sim}{X}$, or equivalently $\underset{\sim}{X} = A^{-1}\underset{\sim}{Y}$, where A is the $n \times n$ matrix with (i, j)th element

$$[A]_{ij} = \begin{cases} 1 - \dfrac{1}{n} & i = j \text{ and } i \neq 1, \\ \dfrac{1}{n} & i = 1 \\ -\dfrac{1}{n} & \text{otherwise} \end{cases}$$

that is, we have a linear transformation, and the Jacobian of the transformation does not depend on any y. Note that

$$\sum_{i=1}^{n}(x_i - \mu)^2 = \sum_{i=1}^{n}(x_i - \overline{x} + \overline{x} - \mu)^2 = \sum_{i=1}^{n}(x_i - \overline{x})^2 + n\left(\overline{x} - \mu\right)^2$$

where $\overline{x} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} x_i$. Note also that the joint pdf of $X_1, ..., X_n$ is, in scalar form

$$\begin{aligned} f_{X_1,..,X_n}(x_1, .., x_n) &= \left(\frac{1}{2\pi\sigma^2}\right)^{n/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\right\} \\ &= \left(\frac{1}{2\pi\sigma^2}\right)^{n/2}\exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(x_i - \overline{x})^2 + n\left(\bar{x} - \mu\right)^2\right]\right\}. \end{aligned}$$

Now

$$x_1 - \overline{x} = -\sum_{i=2}^{n}(x_i - \overline{x}) = -\sum_{i=2}^{n} y_i$$

and so

$$\sum_{i=1}^{n}(x_i-\overline{x})^2=(x_1-\overline{x})^2+\sum_{i=2}^{n}(x_i-\overline{x})^2=\left(-\sum_{i=2}^{n}y_i\right)^2+\sum_{i=2}^{n}y_i^2$$

The Jacobian of the transformation is n, so the joint density of $Y_1,...,Y_n$ is given by direct substitution into (1)

$$\begin{aligned} f_{Y_1,..,Y_n}(y_1,..,y_n) &= n\left(\frac{1}{2\pi\sigma^2}\right)^{n/2}\exp\left\{-\frac{1}{2\sigma^2}\left[\left(-\sum_{i=2}^{n}y_i\right)^2+\sum_{i=2}^{n}y_i^2+n\,(y_1-\mu)^2\right]\right\} \\ &= n\left(\frac{1}{2\pi\sigma^2}\right)^{n/2}\exp\left\{-\frac{1}{2\sigma^2}\left[\left(-\sum_{i=2}^{n}y_i\right)^2+\sum_{i=2}^{n}y_i^2\right]\right\}\times\exp\left\{-\frac{n}{2\sigma^2}\,(y_1-\mu)^2\right\} \end{aligned}$$

Hence

$$f_{Y_1,..,Y_n}(y_1,..,y_n) = f_{Y_2,..,Y_n}(y_2,..,y_n)f_{Y_1}(y_1)$$

and therefore Y_1 is independent of $Y_2,...,Y_n$. Hence $\overline{X}$ is **independent** of the random variables terms $\left\{Y_i=X_i-\bar{X}, i=2,...,n\right\}$. Finally, $\overline{X}$ is also independent of $X_1-\overline{X}$ as

$$X_1-\overline{X}=-\sum_{i=2}^{n}\left(X_i-\overline{X}\right)$$

(b) s^2 is a function only of $\left\{X_i-\overline{X},\ i=1,...,n\right\}$. As $\overline{X}$ is independent of these variables, $\overline{X}$ and s^2 are also independent.

(c)The random variables that appear as sums of squares terms that joint pdf are

$$\frac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\sigma^2}=\frac{\sum\limits_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{\sigma^2}+\frac{n\left(\overline{X}-\mu\right)^2}{\sigma^2}$$

or $V_1=V_2+V_3$, say. Now, $X_i\sim N(\mu,\sigma^2)$, so therefore

$$\frac{(X_i-\mu)^2}{\sigma^2}\sim N(0,1)\Longrightarrow\frac{(X_i-\mu)^2}{\sigma^2}\sim\chi_1^2\equiv Ga\left(\frac{1}{2},\frac{1}{2}\right)\Longrightarrow\frac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\sigma^2}=V_1\sim\chi_n^2$$

as the X_is are independent, and the sum of n independent $Ga(1/2,1/2)$ variables has a $Ga(n/2,1/2)$ distribution. Similarly, as $\overline{X}\sim N(\mu,\sigma^2/n)$, $V_3\sim\chi_1^2$ By part (b), V_2 and V_3 are independent, and so the mgfs of V_1, V_2 and V_3 are related by

$$M_{V_1}(t)=M_{V_2}(t)M_{V_3}(t)\Longrightarrow M_{V_2}(t)=\frac{M_{V_1}(t)}{M_{V_3}(t)}$$

As V_1 and V_3 are Gamma random variables, M_{V_1} and M_{V_3} are given by

$$M_{V_1}(t)=\left(\frac{1/2}{1/2-t}\right)^{n/2},M_{V_3}(t)=\left(\frac{1/2}{1/2-t}\right)^{1/2}\Longrightarrow M_{V_2}(t)=\left(\frac{1/2}{1/2-t}\right)^{(n-1)/2}$$

which is also the mgf of a Gamma random variable, and hence

$$V_2=\frac{(n-1)s^2}{\sigma^2}\sim\chi_{n-1}^2$$