The Multivariate Normal Distribution

The <u>Moment Generating Function</u> (MGF) of a random vector $\boldsymbol{X}$ is given by

$$M_{\boldsymbol{X}}(\boldsymbol{t}) = E(e^{\boldsymbol{t}'\boldsymbol{X}})$$

provided $\exists\ h > 0 \ni E(e^{\boldsymbol{t}'\boldsymbol{X}})$ exists $\forall\ \boldsymbol{t} = (t_1, \ldots, t_n)' \ni t_i \in (-h, h)\ \forall\ i = 1, \ldots, n.$

Result 5.1:

If the MGFs of two random vectors $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ exist in an open rectangle $\mathcal{R}$ that includes the origin, then the cumulative distribution functions (CDFs) of $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are identical iff

$$M_{\boldsymbol{X}_1}(\boldsymbol{t}) = M_{\boldsymbol{X}_2}(\boldsymbol{t}) \quad \forall\, \boldsymbol{t} \in \mathcal{R}.$$

A random variable Z with MGF

$$M_Z(t) = E(e^{tZ}) = e^{t^2/2}$$

is said to have a standard normal distribution.

Show that a random variable Z with density

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

has a standard normal distribution.

$$
\begin{aligned}
E(e^{tZ}) &= \int_{-\infty}^{\infty} e^{tz} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-1/2(z^2-2tz)} dz \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-1/2(z^2-2tz+t^2-t^2)} dz \\
&= e^{t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-1/2(z-t)^2} dz \\
&= e^{t^2/2}.
\end{aligned}
$$

Suppose Z has a standard normal distribution. Then

$$
\begin{aligned}
E(Z) &= \left.\frac{\partial M_Z(t)}{\partial t}\right|_{t=0} \\
&= \left.\frac{\partial e^{t^2/2}}{\partial t}\right|_{t=0} \\
&= e^{t^2/2}(t)|_{t=0} = 0.
\end{aligned}
$$

$$
\begin{aligned}
E(Z^2) &= \left.\frac{\partial^2 M_Z(t)}{\partial t^2}\right|_{t=0} \\
&= e^{t^2/2} + t^2 e^{t^2/2}|_{t=0} \\
&= 1.
\end{aligned}
$$

Thus,

$$E(Z) = 0 \quad \text{and} \quad \text{Var}(Z) = 1.$$

If Z is standard normal, then $Y = \mu + \sigma Z$ has mean

$$E(Y) = E(\mu + \sigma Z) = \mu + \sigma E(Z) = \mu$$

and variance

$$\begin{aligned} \text{Var}(Y) &= \text{Var}(\mu + \sigma Z) \\ &= \text{Var}(\sigma Z) \\ &= \sigma^2 \text{Var}(Z) \\ &= \sigma^2. \end{aligned}$$

Furthermore, the MGF of Y is

$$\begin{aligned} M_Y(t) &= E(e^{tY}) \\ &= E(e^{t(\mu+\sigma Z)}) \\ &= e^{t\mu}E(e^{t\sigma Z}) \\ &= e^{t\mu}M_Z(t\sigma) \\ &= e^{t\mu}e^{t^2\sigma^2/2} \\ &= e^{t\mu+t^2\sigma^2/2}. \end{aligned}$$

A random variable Y with MGF

$$M_Y(t) = e^{t\mu + t^2\sigma^2/2}$$

is said to have a normal distribution with mean μ and variance σ^2.

We denote the distribution of Y by $N(\mu, \sigma^2)$.

If $Y \sim N(\mu, \sigma^2)$, then

$$\begin{aligned}\mathbb{P}(Y \leq y) &= \mathbb{P}(\mu + \sigma Z \leq y) \\ &= \mathbb{P}(Z \leq \frac{y-\mu}{\sigma}).\end{aligned}$$

Thus, the density of Y is

$$\begin{aligned}\frac{\partial \mathbb{P}(Y \leq y)}{\partial y} &= \frac{\partial \mathbb{P}(Z \leq \frac{y-\mu}{\sigma})}{\partial y} \\ &= f_Z\left(\frac{y-\mu}{\sigma}\right)\frac{1}{\sigma} = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2}.\end{aligned}$$

That is,

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2}$$
$$= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(y-\mu)^2}.$$

Suppose $Z_1, \ldots, Z_p \overset{i.i.d.}{\sim} N(0, 1)$.

Then $\boldsymbol{Z} = \begin{bmatrix} Z_1 \\ \vdots \\ Z_p \end{bmatrix}$ is said to have a standard multivariate normal distribution.

$$E(\boldsymbol{Z}) = \boldsymbol{0}$$

$$\text{Var}(\boldsymbol{Z}) = \boldsymbol{I}.$$

Find the Moment Generating Function of a standard multivariate normal random vector $\underset{p\times 1}{\boldsymbol{Z}}$.

$$
\begin{aligned}
E(e^{\boldsymbol{t}'\boldsymbol{Z}}) &= E(e^{\sum_{i=1}^{p} t_i Z_i}) \\
&= E(\prod_{i=1}^{p} e^{t_i Z_i}) \\
&= \prod_{i=1}^{p} E(e^{t_i Z_i}) \\
&= \prod_{i=1}^{p} M_{Z_i}(t_i) \\
&= \prod_{i=1}^{p} e^{t_i^2/2} \\
&= e^{\sum_{i=1}^{p} t_i^2/2} \\
&= e^{\boldsymbol{t}'\boldsymbol{t}/2}.
\end{aligned}
$$

A p-dimensional random vector $\boldsymbol{Y}$ has the <u>Multivariate Normal Distribution</u> with mean $\boldsymbol{\mu}$ and variance $\boldsymbol{\Sigma}$ $(\boldsymbol{Y} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}))$ iff the MGF of $\boldsymbol{Y}$ is

$$M_{\boldsymbol{Y}}(\boldsymbol{t}) = e^{\boldsymbol{t}'\boldsymbol{\mu} + \boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}/2}.$$

Suppose $\boldsymbol{Z} \sim N(\boldsymbol{0}, \boldsymbol{I})$.

Show that

$$\boldsymbol{Y} = \boldsymbol{\mu} + \boldsymbol{A}\boldsymbol{Z}$$

has a multivariate normal distribution.

If $\boldsymbol{Z}$ standard multivariate normal, then $\boldsymbol{Y} = \boldsymbol{\mu} + \boldsymbol{A}\boldsymbol{Z}$ has mean

$$\begin{aligned} E(\boldsymbol{Y}) &= E(\boldsymbol{\mu} + \boldsymbol{A}\boldsymbol{Z}) \\ &= \boldsymbol{\mu} + \boldsymbol{A}E(\boldsymbol{Z}) \\ &= \boldsymbol{\mu} \end{aligned}$$

and variance

$$\begin{aligned} \mathrm{Var}(\boldsymbol{Y}) &= \mathrm{Var}(\boldsymbol{\mu} + \boldsymbol{A}\boldsymbol{Z}) \\ &= \mathrm{Var}(\boldsymbol{A}\boldsymbol{Z}) \\ &= \boldsymbol{A}\mathrm{Var}(\boldsymbol{Z})\boldsymbol{A}' \\ &= \boldsymbol{A}\boldsymbol{A}'. \end{aligned}$$

Furthermore,

$$\begin{aligned}
M_{\boldsymbol{Y}}(\boldsymbol{t}) &= E(e^{\boldsymbol{t}'\boldsymbol{Y}}) \\
&= E(e^{\boldsymbol{t}'(\boldsymbol{\mu}+\boldsymbol{A}\boldsymbol{Z})}) \\
&= e^{\boldsymbol{t}'\boldsymbol{\mu}}E(e^{\boldsymbol{t}'\boldsymbol{A}\boldsymbol{Z}}) \\
&= e^{\boldsymbol{t}'\boldsymbol{\mu}}M_{\boldsymbol{Z}}(\boldsymbol{A}'\boldsymbol{t}) \\
&= e^{\boldsymbol{t}'\boldsymbol{\mu}+\boldsymbol{t}'\boldsymbol{A}\boldsymbol{A}'\boldsymbol{t}/2}.
\end{aligned}$$

Note that if $rank(\underset{q\times p}{\boldsymbol{A}}) < q$, then

$$\begin{aligned}\mathrm{Var}(\boldsymbol{Y}) &= \mathrm{Var}(\boldsymbol{\mu} + \boldsymbol{A}\boldsymbol{Z}) \\ &= \boldsymbol{A}\boldsymbol{A}'\end{aligned}$$

will be singular.

In this case, the support of the $q \times 1$ random vector $\boldsymbol{Y}$ will lie within a $rank(\boldsymbol{A})(< q)$- dimensional vector space.

Give a specific example of a singular multivariate normal distribution $(\boldsymbol{Y} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \boldsymbol{\Sigma}$ singular$)$.

Suppose $\boldsymbol{Z} = Z \sim N(0, 1)$.

Suppose $\boldsymbol{A} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Then

$$rank(\underset{2\times 1}{\boldsymbol{A}}) = 1 < 2.$$

Let $\boldsymbol{Y} = \boldsymbol{AZ}$. Then

$$\boldsymbol{Y} \sim N\left(\boldsymbol{0}, \boldsymbol{AA}' = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right).$$

$\boldsymbol{Y}$ lies in the 1-dimensional vector space

$\mathcal{C} = \{\boldsymbol{Y} = (Y_1, Y_2)' \in \mathbb{R}^2 : Y_1 = Y_2\}$ with probability 1.

Result 5.3:

If $\boldsymbol{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\boldsymbol{Y} = \boldsymbol{a} + \boldsymbol{B}\boldsymbol{X}$, then

$$\boldsymbol{Y} \sim N(\boldsymbol{a} + \boldsymbol{B}\boldsymbol{\mu}, \boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}').$$

Proof of Result 5.3:

$$
\begin{aligned}
E(e^{\boldsymbol{t}'\boldsymbol{Y}}) &= E(e^{\boldsymbol{t}'\boldsymbol{a}+\boldsymbol{t}'\boldsymbol{B}\boldsymbol{X}}) \\
&= e^{\boldsymbol{t}'\boldsymbol{a}} E(e^{\boldsymbol{t}'\boldsymbol{B}\boldsymbol{X}}) \\
&= e^{\boldsymbol{t}'\boldsymbol{a}} M_{\boldsymbol{X}}(\boldsymbol{B}'\boldsymbol{t}) \\
&= e^{\boldsymbol{t}'\boldsymbol{a}} e^{\boldsymbol{t}'\boldsymbol{B}\mu+\boldsymbol{t}'\boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}'\boldsymbol{t}/2} \\
&= e^{\boldsymbol{t}'(\boldsymbol{a}+\boldsymbol{B}\boldsymbol{\mu})+\boldsymbol{t}'\boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}'\boldsymbol{t}/2}.
\end{aligned}
$$

Thus, $\boldsymbol{Y} \sim N(\boldsymbol{a}+\boldsymbol{B}\boldsymbol{\mu}, \boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}')$. $\square$

Corollary 5.1:

If $\underset{p\times 1}{\boldsymbol{X}}$ is multivariate normal (MVN), then the joint distribution of any subvector of $\boldsymbol{X}$ is MVN.

Proof of Corollary 5.1:

Suppose $\{i_1, \ldots, i_q\} \subseteq \{1, \ldots, p\}$. Then

$$\begin{bmatrix} \boldsymbol{X}_{i_1} \\ \vdots \\ \boldsymbol{X}_{i_q} \end{bmatrix} = \begin{bmatrix} \boldsymbol{e}'_{i_1} \\ \vdots \\ \boldsymbol{e}'_{i_q} \end{bmatrix} \boldsymbol{X},$$

where $\boldsymbol{e}'_i$ denotes the i^{th} row of the $p \times p$ identity matrix.

$\begin{bmatrix} \boldsymbol{e}'_{i_1} \\ \vdots \\ \boldsymbol{e}'_{i_q} \end{bmatrix} \boldsymbol{X} \sim$ MVN by Result 5.3. $\square$

Corollary 5.2:

If $\underset{p\times 1}{\boldsymbol{X}} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\boldsymbol{\Sigma}$ is nonsingular, then

(a) $\exists$ a nonsingular matrix $\boldsymbol{A} \ni \boldsymbol{\Sigma} = \boldsymbol{A}\boldsymbol{A}'$,

(b) $\boldsymbol{A}^{-1}(\boldsymbol{X} - \boldsymbol{\mu}) \sim N(\boldsymbol{0}, \boldsymbol{I})$, and

(c) The probability density function of $\boldsymbol{X}$ is

$$f_{\boldsymbol{X}}(\boldsymbol{t}) = (2\pi)^{-p/2}|\boldsymbol{\Sigma}|^{-1/2}e^{-1/2(\boldsymbol{t}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{t}-\boldsymbol{\mu})}.$$

Proof of Corollary 5.2:

(a) We can take $\boldsymbol{A} = \boldsymbol{\Sigma}^{1/2}$ because $\boldsymbol{\Sigma}$ is symmetric and positive definite. Because $\boldsymbol{\Sigma}$ positive definite, $(\boldsymbol{\Sigma}^{1/2})^{-1} = \boldsymbol{\Sigma}^{-1/2}$ exists.

(b) By Result 5.3,

$$\boldsymbol{A}^{-1}(\boldsymbol{X} - \boldsymbol{\mu}) \sim N(\boldsymbol{A}^{-1}\boldsymbol{\mu} - \boldsymbol{A}^{-1}\boldsymbol{\mu}, \boldsymbol{A}^{-1}\boldsymbol{\Sigma}(\boldsymbol{A}^{-1})'),$$

with

$$\boldsymbol{A}^{-1}\boldsymbol{\mu} - \boldsymbol{A}^{-1}\boldsymbol{\mu} = \boldsymbol{0} \quad \text{and} \quad \boldsymbol{A}^{-1}\boldsymbol{\Sigma}(\boldsymbol{A}^{-1})' = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1/2} = \boldsymbol{I}.$$

(c) Homework problem.

You may wish to use the multivariate change of variables result on page 185 of Casella and Berger. □

Result 5.2:

Suppose the MGF of $\boldsymbol{X}_i$ is $M_{\boldsymbol{X}_i}(\boldsymbol{t}_i)\ \forall\ i = 1, \ldots, p$. Let

$$\boldsymbol{X} = [\boldsymbol{X}_1', \boldsymbol{X}_2', \ldots, \boldsymbol{X}_p']' \quad \text{and} \quad \boldsymbol{t} = [\boldsymbol{t}_1', \boldsymbol{t}_2', \ldots, \boldsymbol{t}_p']'.$$

Suppose $\boldsymbol{X}$ has MGF $M_{\boldsymbol{X}}(\boldsymbol{t})$.

Then $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_p$ are mutually independent iff

$$M_{\boldsymbol{X}}(\boldsymbol{t}) = \prod_{i=1}^{p} M_{\boldsymbol{X}_i}(\boldsymbol{t}_i)$$

$\forall\ \boldsymbol{t}$ in an open rectangle that includes $\mathbf{0}$.

Result 5.4:

If $\boldsymbol{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and we partition

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_1 \\ \vdots \\ \boldsymbol{X}_p \end{bmatrix}, \quad \boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \vdots \\ \boldsymbol{\mu}_p \end{bmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \cdots & \boldsymbol{\Sigma}_{1p} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Sigma}_{p1} & \cdots & \boldsymbol{\Sigma}_{pp} \end{bmatrix}.$$

Then $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_p$ are mutually independent iff

$$\boldsymbol{\Sigma}_{ij} = \boldsymbol{0} \quad \forall\ i \neq j.$$

Proof of Result 5.4:

$(\Longrightarrow)$ $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_p$ mutually independent, then

$$\begin{aligned}
\text{Cov}(\boldsymbol{X}_i, \boldsymbol{X}_j) &= E[(\boldsymbol{X}_i - \boldsymbol{\mu}_i)(\boldsymbol{X}_j - \boldsymbol{\mu}_j)'] \\
&= [E(\boldsymbol{X}_i - \boldsymbol{\mu}_i)][E(\boldsymbol{X}_j - \boldsymbol{\mu}_j)'] \\
&= [\boldsymbol{\mu}_i - \boldsymbol{\mu}_i][(\boldsymbol{\mu}_j - \boldsymbol{\mu}_j)'] \\
&= \boldsymbol{0} \qquad \Rightarrow \quad \forall\ i \neq j.
\end{aligned}$$

Thus, $\boldsymbol{\Sigma}_{ij} = \boldsymbol{0} \quad \forall\ i \neq j$.

($\Longleftarrow$) Partition $\boldsymbol{t}$ as $[\boldsymbol{t}_1', \ldots, \boldsymbol{t}_p']'$.

Then

$$\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t} = \sum_{i=1}^{p}\sum_{j=1}^{p} \boldsymbol{t}_i'\boldsymbol{\Sigma}_{ij}\boldsymbol{t}_j.$$

If $\boldsymbol{\Sigma}_{ij} = \mathbf{0}\ \forall\ i \neq j$, then

$$\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t} = \sum_{i=1}^{p} \boldsymbol{t}_i'\boldsymbol{\Sigma}_{ii}\boldsymbol{t}_i.$$

Thus,

$$\begin{aligned} M_{\boldsymbol{X}}(\boldsymbol{t}) &= e^{\boldsymbol{t}'\boldsymbol{\mu}+\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}/2} = e^{\sum_{i=1}^{p}(\boldsymbol{t}_i'\boldsymbol{\mu}_i+\boldsymbol{t}_i'\boldsymbol{\Sigma}_{ii}\boldsymbol{t}_i/2)} \\ &= \prod_{i=1}^{p} e^{\boldsymbol{t}_i'\boldsymbol{\mu}_i+\boldsymbol{t}_i'\boldsymbol{\Sigma}_{ii}\boldsymbol{t}_i/2} = \prod_{i=1}^{p} M_{\boldsymbol{X}_i}(\boldsymbol{t}_i). \end{aligned}$$

By Result 5.2, $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_p$ are mutually independent. $\square$

Corollary 5.3:

Suppose

$$\boldsymbol{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$
$$\boldsymbol{Y}_1 = \boldsymbol{a}_1 + \boldsymbol{B}_1\boldsymbol{X}, \quad \text{and}$$
$$\boldsymbol{Y}_2 = \boldsymbol{a}_2 + \boldsymbol{B}_2\boldsymbol{X}.$$

Then $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ are independent iff

$$\boldsymbol{B}_1\boldsymbol{\Sigma}\boldsymbol{B}_2' = \boldsymbol{0}.$$

Proof of Corollary 5.3:

Let

$$\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{Y}_1 \\ \boldsymbol{Y}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \end{bmatrix} + \begin{bmatrix} \boldsymbol{B}_1 \\ \boldsymbol{B}_2 \end{bmatrix} \boldsymbol{X}.$$

Then $\boldsymbol{Y}$ MVN with

$$\begin{aligned} \text{Var}(\boldsymbol{Y}) &= \begin{bmatrix} \boldsymbol{B}_1 \\ \boldsymbol{B}_2 \end{bmatrix} \boldsymbol{\Sigma} \begin{bmatrix} \boldsymbol{B}_1' & \boldsymbol{B}_2' \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{B}_1\boldsymbol{\Sigma}\boldsymbol{B}_1' & \boldsymbol{B}_1\boldsymbol{\Sigma}\boldsymbol{B}_2' \\ \boldsymbol{B}_2\boldsymbol{\Sigma}\boldsymbol{B}_1' & \boldsymbol{B}_2\boldsymbol{\Sigma}\boldsymbol{B}_2' \end{bmatrix}. \end{aligned}$$

By Result 5.4, $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ independent $\iff \boldsymbol{B}_1\boldsymbol{\Sigma}\boldsymbol{B}_2' = \boldsymbol{0}$. $\square$