

5 Multivariate Normal Distribution

- Density (Let $\mathbf{Y}_{p \times 1} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$) $\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix}$

$$f_{\mathbf{Y}}(\mathbf{y}) = |\boldsymbol{\Sigma}|^{-\frac{1}{2}} (2\pi)^{-\frac{p}{2}} e^{-\frac{1}{2} \{(\mathbf{y} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \boldsymbol{\mu})\}}$$

- MGF: $M_{\mathbf{Y}}(\mathbf{t}) = e^{\mathbf{t}'\boldsymbol{\mu} + \frac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}}$ - Remark 5.1

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \sim N \left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \right)$$

$$Y_1 \sim N(\mu_1, b_{11}) \quad b_{11} = b_1^2$$

$$Y_2 \sim N(\mu_2, b_{22}) \quad b_{22} = b_2^2 \quad \boldsymbol{\Sigma}: \text{Covariance matrix}$$

Remark 5.1

$$M_Y(t) = E_X(e^{t'X})$$

$$= |\Sigma|^{-\frac{1}{2}} (2\pi)^{-\frac{p}{2}} \int_{\mathbb{R}^p} e^{t'y} e^{-\frac{1}{2}(y-\mu)'\Sigma^{-1}(y-\mu)} dy$$

$$= |\Sigma|^{-\frac{1}{2}} (2\pi)^{-\frac{p}{2}} \int e^{-\frac{1}{2}[(y-\mu)'\Sigma^{-1}(y-\mu) - 2t'y]} dy$$

$$(y-\mu)'\Sigma^{-1}(y-\mu) - 2t'y$$

$$(y-\mu)'\Sigma^{-1}(y-\mu)$$

$$= y'\Sigma^{-1}y - 2\mu'\Sigma^{-1}y + \mu'\Sigma^{-1}\mu - 2t'y$$

$$= (y'\Sigma^{-1} - \mu'\Sigma^{-1})(y-\mu)$$

$$= y'\Sigma^{-1}y - 2(\Sigma^{-1}\mu + t)'y + \mu'\Sigma^{-1}\mu$$

$$= y'\Sigma^{-1}y - \mu'\Sigma^{-1}y - y'\Sigma^{-1}\mu + \mu'\Sigma^{-1}\mu$$

$$\Sigma^{-1}\mu + t \stackrel{!}{=} \mu^*$$

$$\mu'\Sigma^{-1}y = y'\Sigma^{-1}\mu$$

$$= (y-\mu^*)'\Sigma^{-1}(y-\mu^*) - \mu^{*'}\Sigma^{-1}\mu^* + \mu'\Sigma^{-1}\mu$$

$$= (y-\mu^*)'\Sigma^{-1}(y-\mu^*) - t'\Sigma^{-1}t - 2t'\mu \quad \dots (*)$$

$$= |\Sigma|^{-\frac{1}{2}} (2\pi)^{-\frac{p}{2}} \int e^{-\frac{1}{2}(y-\mu^*)'\Sigma^{-1}(y-\mu^*)} dy e^{t'\mu + \frac{1}{2}t'\Sigma^{-1}t}$$

pdf of $N(\mu^*, \Sigma)$

$$= \exp\left[t'\mu + \frac{1}{2}t'\Sigma^{-1}t\right]$$

1°: If A n.s. $\rightarrow A'$ n.s.

$$(A')' = (A')'$$

$$2^\circ: (A')^T = (A^T)' = A'$$

- Let \mathbf{B} be a constant matrix and \mathbf{C} be a constant vector

Remark 5.2 $\mathbf{BY} + \mathbf{C} \sim N(\mathbf{B}\boldsymbol{\mu} + \mathbf{C}, \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}')$

- Marginal Distribution, Condition Distribution and independence

Let

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{bmatrix} \sim N \left[\begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix} \right]$$

then

(i) $\mathbf{Y}_1 \sim N(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$

(ii) $\mathbf{Y}_1 | \mathbf{Y}_2 = \mathbf{y}_2 \sim N(\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{y}_2 - \boldsymbol{\mu}_2), \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21})$

(iii) \mathbf{Y}_1 and \mathbf{Y}_2 are independent iff $\boldsymbol{\Sigma}_{12} = \mathbf{0}$

Remark 5.3

Remark 5.2

1°. $\underline{B}\underline{X} + \underline{c} \sim N(\quad), \text{ if } \underline{X} \sim N(\quad).$

2°. $E(\underline{B}\underline{X} + \underline{c}) = \underline{B}\underline{\mu} + \underline{c}$

3°. $\text{Var}(\underline{B}\underline{X} + \underline{c}) = \text{Var}(\underline{B}\underline{X}) = \underline{B} \text{Var}(\underline{X}) \underline{B}' = \underline{B}\underline{\Sigma}\underline{B}'$

4°. Proof: use MGF

$$\begin{aligned}
 M_{\underline{B}\underline{X} + \underline{c}}(t) &= E(e^{t'(\underline{B}\underline{X} + \underline{c})}) \\
 &= E(e^{t'\underline{B}\underline{X} + t'\underline{c}}) \\
 &= E(e^{t^*\underline{X}}) \cdot e^{t'\underline{c}} \quad \underline{t}'\underline{B} \triangleq \underline{t}^* \\
 &= \exp\left\{ \underline{t}^*\underline{\mu} + \frac{1}{2} \underline{t}^* \underline{\Sigma} \underline{t}^* \right\} \cdot \exp\{t'\underline{c}\} \\
 &= \exp\left\{ t'\underline{B}\underline{\mu} + \frac{1}{2} t'\underline{B}\underline{\Sigma}\underline{B}'t + t'\underline{c} \right\} \\
 &= e^{t'(\underline{B}\underline{\mu} + \underline{c}) + \frac{1}{2} t'\underline{B}\underline{\Sigma}\underline{B}'t} \\
 &= e^{t'\underline{\mu}^* + \frac{1}{2} t'\underline{\Sigma}^*t} \quad \underline{\mu}^* \triangleq \underline{B}\underline{\mu} + \underline{c} \\
 &\sim N(\underline{\mu}^*, \underline{\Sigma}^*) \quad \underline{\Sigma}^* \triangleq \underline{B}\underline{\Sigma}\underline{B}'
 \end{aligned}$$

Remark 5.3

$P(\underline{y}_1 | \underline{y}_2) = \frac{P(\underline{y}_1, \underline{y}_2)}{P(\underline{y}_2)} \propto \exp\left\{ -\frac{1}{2} \begin{pmatrix} \underline{y}_1 - \underline{\mu}_1 \\ \underline{y}_2 - \underline{\mu}_2 \end{pmatrix}' \underline{\Sigma}^{-1} \begin{pmatrix} \underline{y}_1 - \underline{\mu}_1 \\ \underline{y}_2 - \underline{\mu}_2 \end{pmatrix} \right\}$ 偷懒x1 关心y1, 其余可以当作 constant.

$\Rightarrow \underline{\Sigma}^{-1} = \begin{pmatrix} \underline{\Sigma}_{11} & \underline{\Sigma}_{12} \\ \underline{\Sigma}_{21} & \underline{\Sigma}_{22} \end{pmatrix} = \begin{pmatrix} \underline{B}^{-1} & -\underline{B}^{-1}\underline{\Sigma}_{22}^{-1}\underline{\Sigma}_{21} \\ -\underline{\Sigma}_{22}^{-1}\underline{\Sigma}_{21}\underline{B}^{-1} & * \end{pmatrix}$ where $\underline{B}^{-1} = \underline{\Sigma}_{11} - \underline{\Sigma}_{12}\underline{\Sigma}_{22}^{-1}\underline{\Sigma}_{21}$

$$\begin{pmatrix} y_1 - \mu_1 \\ y_2 - \mu_2 \end{pmatrix}' \tilde{\Sigma}^{-1} \begin{pmatrix} y_1 - \mu_1 \\ y_2 - \mu_2 \end{pmatrix} = (y_1 - \mu_1)' B^{-1} (y_1 - \mu_1) - 2(y_1 - \mu_1)' \tilde{B}^{-1} \tilde{\Sigma}_{12} \tilde{\Sigma}_{22}^{-1} (y_2 - \mu_2) + C \quad \text{偷懒} \times 2.$$

$$\text{跳跃} \quad = (y_1 - \mu_1)' B^{-1} (y_1 - \mu_1) - 2(y_1 - \mu_1)' \tilde{B}^{-1} \mu^* + C$$

$$= (y_1 - \mu_1 - \mu^*)' B^{-1} (y_1 - \mu_1 - \mu^*) + C^*$$

$$= \tilde{y}_1^*{}' \tilde{B}^{-1} \tilde{y}_1^* - 2 \tilde{y}_1^*{}' \tilde{B}^{-1} \mu^* + C^*$$

$$= (\tilde{y}_1^* - \mu^*)' \tilde{B}^{-1} (\tilde{y}_1^* - \mu^*) + C^*$$

$$\mu^* = \tilde{\Sigma}_{12} \tilde{\Sigma}_{22}^{-1} (y_2 - \mu_2)$$

$$y_1^* = y_1 - \mu_1$$

$$P(y_1 | y_2) \propto \exp \left\{ -\frac{1}{2} (y_1 - \mu_1 - \mu^*)' \tilde{B}^{-1} (y_1 - \mu_1 - \mu^*) \right\}$$

$$\sim N(\mu_1 + \mu^*, B) = N(\mu_1 + \tilde{\Sigma}_{12} \tilde{\Sigma}_{22}^{-1} (y_2 - \mu_2), \tilde{\Sigma}_{11} - \tilde{\Sigma}_{12} \tilde{\Sigma}_{22}^{-1} \tilde{\Sigma}_{21})$$

$$= (y_1 - \mu_1)' \tilde{B}^{-1} (y_1 - \mu_1) - 2(y_1 - \mu_1)' \tilde{B}^{-1} \mu^* + C$$

understand: 当给定你爸身高时,你爷爷的身高的信息量几乎包含在你爸提供的信息中?

- Partial Correlation Conditional Correlation

Let $\mathbf{v} \sim N_q(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and

$$\mathbf{v} = \begin{pmatrix} \mathbf{y} \\ \mathbf{x} \end{pmatrix}; \quad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_y \\ \boldsymbol{\mu}_x \end{pmatrix}; \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{yy} & \boldsymbol{\Sigma}_{yx} \\ \boldsymbol{\Sigma}_{xy} & \boldsymbol{\Sigma}_{xx} \end{pmatrix}$$

where $\mathbf{y} = (y_1, y_2, \dots, y_{r-1})'$ and $\mathbf{x} = (x_r, \dots, x_q)'$. Let $\rho_{ij.r\dots q}$ be the partial correlation between y_i and y_j , $1 \leq i < j \leq r-1$, in the conditional distribution of \mathbf{y} given \mathbf{x} . By the definition of correlation, we have

Remark 5.4

$$\rho_{ij.r\dots q} = \frac{\sigma_{ij.r\dots q}}{\sqrt{\sigma_{ii.r\dots q}\sigma_{jj.r\dots q}}}.$$

Matrix of partial correlations

$$\boldsymbol{\Omega}_{y.x} = \mathbf{D}_{y.x}^{-1} \boldsymbol{\Sigma}_{y.x} \mathbf{D}_{y.x}^{-1}$$

where $\boldsymbol{\Sigma}_{y.x} = \boldsymbol{\Sigma}_{yy} - \boldsymbol{\Sigma}_{yx} \boldsymbol{\Sigma}_{xx}^{-1} \boldsymbol{\Sigma}_{xy}$ and $\mathbf{D}_{y.x} = [\text{diag}(\boldsymbol{\Sigma}_{y.x})]^{1/2}$.

$$\boldsymbol{\Omega}_{y.x} = \mathbf{D}_{y.x}^{-1} \boldsymbol{\Sigma}_{y.x} \mathbf{D}_{y.x}^{-1}$$

Remark 5.4

Correlation and partial correlation

Y — blood pressure.

X_1 — amount of food eaten

X_2 — weight

$$\rho_{Y, X_1} = \text{Corr}(Y, X_1)$$

In general, $\text{Corr}(X_1, X_2) \neq 0$. $\text{Corr}(Y, X_2)$ — need to consider confounder X_1 .

— If we need to add X_2 in the model.

We need to consider partial correlation between Y and X_2 . given X_1

— how to do?

$$Y = a + bX_1 + \varepsilon, \quad \begin{matrix} \text{residual} \\ \Rightarrow r_{Y, X_1} = Y - (\hat{a} + \hat{b}X_1) \end{matrix}$$

$$X_2 = c + dX_1 + \varepsilon_1 \Rightarrow r_{X_2, X_1} = X_2 - (\hat{c} + \hat{d}X_1)$$

$$\text{Corr}(r_{Y, X_1}, r_{X_2, X_1}) = \rho_{(Y, X_2), X_1} \quad \text{partial correlation between } Y \text{ and } X_2 \text{ given } X_1$$

— if $\rho_{(Y, X_2), X_1} = 0$, all the information in X_2 to explain the variation of Y is included in X_1 .

We don't need to add X_2 to the model if X_1 is already used in the model.

Ch5. Remark 5.4

$$y = \alpha + b x_1 + \varepsilon \Rightarrow r_{y, x_1} = y - (\hat{\alpha} + \hat{b} x_1)$$

$$x_2 = c + d x_1 + \varepsilon \Rightarrow r_{x_2, x_1} = x_2 - (\hat{c} + \hat{d} x_1)$$

$\text{Corr}(r_{y, x_1}, r_{x_2, x_1}) = \rho_{(y, x_2), x_1}$ partial correlation between y and x_2 given x_1

$$\begin{pmatrix} y \\ x_1 \\ x_2 \end{pmatrix} \sim N \left(\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix}, \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} \right)$$

从上一节课

$$\begin{pmatrix} y \\ x_2 \end{pmatrix} | x_1 \sim N \left(\mu_{(y, x_2) | x_1}, \Sigma_{(y, x_2) | x_1} \right)$$

用线性模型解释一下, 方便理解.

$$\text{where } \mu_{(y, x_2) | x_1} = \begin{pmatrix} \mu_{y | x_1} \\ \mu_{x_2 | x_1} \end{pmatrix}, \Sigma_{(y, x_2) | x_1} = \begin{pmatrix} b_{yy | x_1} & b_{yx_2 | x_1} \\ b_{x_2 y | x_1} & b_{x_2 x_2 | x_1} \end{pmatrix}$$

$$\Rightarrow \frac{b_{yx_2 | x_1}}{\sqrt{b_{yy | x_1} \cdot b_{x_2 x_2 | x_1}}} = \rho_{(y, x_2), x_1} \quad \text{可以证明}$$

In general.

$$\mathcal{Y} = \begin{pmatrix} y \\ \mathbf{x} \end{pmatrix}^{r\text{-dim}} \quad \mathcal{Y} \sim N(\mu, \Sigma), \mu = \begin{pmatrix} \mu_y \\ \mu_x \end{pmatrix}, \Sigma = \begin{pmatrix} \Sigma_{yy} & \Sigma_{yx} \\ \Sigma_{xy} & \Sigma_{xx} \end{pmatrix}$$

$$\text{Cov}(\mathcal{Y} | \mathbf{x}) = \Sigma_{yy} - \Sigma_{yx} \Sigma_{xx}^{-1} \Sigma_{xy}$$

$$= \begin{pmatrix} b_{11|x} & b_{12|x} & \dots & b_{1n|x} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1|x} & b_{n2|x} & \dots & b_{nn|x} \end{pmatrix}_{r \times r}$$

jj: x
means: given

Define $\rho_{ij|x} = \frac{b_{ij|x}}{\sqrt{b_{ii|x}} \sqrt{b_{jj|x}}}$ partial correlation between y_i and y_j given \mathbf{x}

partial correlation matrix:

$$\rho_{\mathcal{Y}|\mathbf{x}} = \begin{pmatrix} 1 & \rho_{12|x} & \dots & \rho_{1n|x} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1|x} & \rho_{n2|x} & \dots & 1 \end{pmatrix}$$

$$\begin{pmatrix} \frac{33}{14} & \frac{1}{14} & -\frac{9}{14} \\ \frac{1}{14} & \frac{3}{14} & \frac{1}{14} \\ -\frac{9}{14} & \frac{1}{14} & \frac{5}{14} \end{pmatrix}$$

$$\Rightarrow \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}$$

$$= 9 - (0, 3, 3) \begin{pmatrix} & & \\ & & \\ & & \end{pmatrix} \begin{pmatrix} 0 \\ 3 \\ 7 \end{pmatrix}$$

Example 5.1

$$\mathbf{y} = \begin{pmatrix} y \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} \sim N \left(\begin{pmatrix} 2 \\ 5 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 9 & 0 & 3 & 3 \\ 0 & 1 & -1 & 2 \\ 3 & -1 & 6 & -3 \\ 3 & 2 & -3 & 7 \end{pmatrix} \right) \leftarrow \text{sample covariance}$$

$$\begin{pmatrix} \mu_y \\ \mu_x \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \\ -2 \\ 1 \end{pmatrix} \quad \Sigma = \begin{pmatrix} b_{yy} & b'_{yx} \\ b_{yx} & \Sigma_{xx} \end{pmatrix}$$

$$y|x \sim N(E(y|x), \text{Var}(y|x))$$

$$E(y|x) = \mu_y + b'_{yx} \Sigma_{xx}^{-1} (x - \mu_x)$$

$$= 2 + (0, 3, 3) \begin{pmatrix} -1 & -1 & 2 \\ -1 & 6 & -3 \\ 2 & -3 & 7 \end{pmatrix}^{-1} \begin{pmatrix} x_1 - 5 \\ x_2 + 2 \\ x_3 - 1 \end{pmatrix}$$

$$= \dots$$

$$= \frac{95}{7} - \frac{12}{7} x_1 + \frac{6}{7} x_2 + \frac{9}{7} x_3$$

$$\text{Var}(y|x) = b_{yy} - b'_{yx} \Sigma_{xx}^{-1} b_{yx}$$

$$= 9 - (0, 3, 3) \begin{pmatrix} & & \\ & & \\ & & \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 3 \\ 3 \end{pmatrix} = \frac{18}{7} \leftarrow \checkmark$$

Remarks for ex. 5.1

$$E(y|x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

$$(\beta_0, \beta_1, \beta_2, \beta_3) = \left(\frac{95}{7}, -\frac{12}{7}, \frac{6}{7}, \frac{9}{7} \right)$$

Suppose (y, x) 's obey multivariate normal distribution

\Rightarrow Conditional mean of $y|x$ is linear regression model

conditional distribution

\downarrow
con. mean and con. variance

$$2 + \left(-\frac{12}{7} \frac{6}{7} \frac{9}{7} \right)$$

$$2 + -\frac{12}{7} (x_1 - 5)$$

$$+ \frac{6}{7} (x_2 + 2)$$

$$+ \frac{9}{7} (x_3 - 1)$$

$$2 - \frac{12}{7} x_1 + \frac{60}{7} + \frac{12}{7} - \frac{9}{7} \frac{63}{7} + 7$$

$$\frac{63}{7} + 2 \quad 14 \quad \frac{75}{7}$$

Example 5.2

$$\mathbf{y} = \begin{pmatrix} y \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} \sim N \left(\begin{pmatrix} 2 \\ 5 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 9 & 0 & 3 & 3 \\ 0 & 1 & -1 & 2 \\ 3 & -1 & 6 & -3 \\ 3 & 2 & -3 & 7 \end{pmatrix} \right) \leftarrow \text{sample covariance}$$

$$42 - (9) \quad 33$$

$n \rightarrow \infty, S \rightarrow \Sigma$

$$\begin{pmatrix} \mu_y \\ \mu_{x_1} \\ \mu_{x_2} \\ \mu_{x_3} \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \\ -2 \\ 1 \end{pmatrix} \quad \Sigma = \begin{pmatrix} \sigma_{yy} & \sigma_{yx_1} \\ \sigma_{yx_1} & \sigma_{x_1x_1} \end{pmatrix}$$

$$\text{Cov}(y|x) = \Sigma_{yy} - \Sigma_{yx} \Sigma_{xx}^{-1} \Sigma_{xy}$$

$$= \begin{pmatrix} 9 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 3 & 3 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 6 & -3 \\ -3 & 7 \end{pmatrix}^{-1} \begin{pmatrix} 3 & -1 \\ 3 & 2 \end{pmatrix}$$

$$= \frac{1}{33} \begin{pmatrix} 126 & -24 \\ -24 & 14 \end{pmatrix}$$

$$\rho_{y|x_1, (x_2, x_3)} = \frac{-24/33}{\sqrt{126/33} \cdot \sqrt{14/33}} = -.571$$

$$\rho_{yx_1} = 0$$

sub-group analysis.

mediation analysis



$$\begin{pmatrix} 9 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 3 & 3 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 6 & -3 \\ -3 & 7 \end{pmatrix}^{-1} \begin{pmatrix} 3 & -1 \\ 3 & 2 \end{pmatrix}$$

$$= \begin{pmatrix} 9 & 0 \\ 0 & 1 \end{pmatrix} - \frac{1}{33} \begin{pmatrix} 3 & 3 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 7 & 3 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} 3 & -1 \\ 3 & 2 \end{pmatrix}$$

$$= \begin{pmatrix} 9 & 0 \\ 0 & 1 \end{pmatrix} - \frac{1}{33} \begin{pmatrix} 30 & 27 \\ -1 & 9 \end{pmatrix} \begin{pmatrix} 3 & -1 \\ 3 & 2 \end{pmatrix}$$



如何解释

如何解释