Multivariate Normal

Result 4.6:

$$\underset{p\times 1}{X} = \begin{bmatrix} X_1 \\ \hdashline X_2 \end{bmatrix} \sim N_p(\mu, \Sigma), \quad \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}, \quad |\Sigma_{22}| > 0$$

Then

$$X_1 \mid X_2 = x_2 \sim N(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2),\ \Sigma_{11\cdot 2})$$

where $\Sigma_{11\cdot 2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$

An indirect proof:

Start from the basic fact that:

Fact 1: if A and D are square matrices:

$$\begin{vmatrix} A & B \\ 0 & D \end{vmatrix} = |A|\,|D|$$

Idea: If $X_1 \perp X_2$, then $X_1 \mid X_2 = x_2$ has the same distribution with X_1.

Fact 2: For normal variables, independent $\Leftrightarrow$ uncorrelated.

$$\begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \xrightarrow{\text{Step 1}} \left[\begin{array}{c:c} \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} & 0' \\ \hdashline \Sigma_{21} & \Sigma_{22} \end{array}\right] \xrightarrow{\text{Step 2.}} \left[\begin{array}{c:c} \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} & 0' \\ \hdashline 0 & \Sigma_{22} \end{array}\right]$$

Step 1:

$$\underbrace{\left[\begin{array}{c:c} I_{q\times q} & -\Sigma_{12}\Sigma_{22}^{-1} \\ \hdashline 0_{(p-q)\times q} & I_{(p-q)\times(p-q)} \end{array}\right]}_{A} \underbrace{\left[\begin{array}{c:c} \Sigma_{11} & \Sigma_{12} \\ \hdashline \Sigma_{21} & \Sigma_{22} \end{array}\right]}_{A'}$$

Step 2: $\left[\begin{array}{c|c} I_{q\times q} & -\Sigma_{12}\Sigma_{22}^{-1} \\ \hline 0_{(p-q)\times q} & I_{(p-q)\times(p-q)} \end{array}\right] \left[\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{21} & \Sigma_{22} \end{array}\right] \left[\begin{array}{c|c} I_{q\times q} & 0_{q\times(p-q)} \\ \hline (-\Sigma_{12}\Sigma_{22}^{-1})' & I_{(p-q)\times(p-q)} \end{array}\right]$

Denote $A = \begin{bmatrix} I & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I \end{bmatrix}$

$\Rightarrow \boxed{A\Sigma A'} = \begin{pmatrix} \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} & 0' \\ 0 & \Sigma_{22} \end{pmatrix}$

Cov of $A\underline{X}$ cov of $\underline{X}$

Hence, if we do the linear transformation:

$$\underline{Y} = A\underline{X}, \quad \underline{Y} \sim N(A\underline{\mu}, A\Sigma A')$$

Moreover, we can centralize it:

$$\underline{Y} = A(\underline{X} - \underline{\mu}) = \begin{bmatrix} \underline{X}_1 - \underline{\mu}_1 - \Sigma_{12}\Sigma_{22}^{-1}(\underline{X}_2 - \underline{\mu}_2) \\ \underline{X}_2 - \underline{\mu}_2 \end{bmatrix}$$

$$= \begin{bmatrix} \underline{Y}_1 \\ \underline{Y}_2 \end{bmatrix}$$

$$\underline{Y}_1 \sim N(\underline{0}, \Sigma_{11\cdot 2}), \quad \underline{Y}_2 \sim N(\underline{0}, \Sigma_{22})$$

$$\underline{Y}_1 \mid \underline{Y}_2 \sim N(\underline{0}, \Sigma_{11\cdot 2}) \quad \text{(Independence)}$$

$$\Rightarrow \underline{Y}_1 \mid \underline{X}_2 \sim N(\underline{0}, \Sigma_{11\cdot 2})$$

$$\Rightarrow \underline{X}_1 = \underline{Y}_1 + \underline{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\underline{X}_2 - \underline{\mu}_2) \mid \underline{X}_2 \sim N(\underline{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\underline{X}_2 - \underline{\mu}_2), \Sigma_{11\cdot 2})$$

Method 2 : Direct proof. (Ex. 4.13)

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

$$= \begin{bmatrix} \underline{x}_1-\underline{\mu}_1 \\ \underline{x}_2-\underline{\mu}_2 \end{bmatrix}' \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}^{-1} \begin{bmatrix} \underline{x}_1-\underline{\mu}_1 \\ \underline{x}_2-\underline{\mu}_2 \end{bmatrix}$$

$$= \begin{bmatrix} \underline{x}_1-\underline{\mu}_1 \\ \underline{x}_2-\underline{\mu}_2 \end{bmatrix}' \begin{bmatrix} I & \underline{0} \\ -\Sigma_{22}^{-1}\Sigma_{21} & I \end{bmatrix} \begin{bmatrix} (\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1} & \underline{0} \\ \underline{0}' & \Sigma_{22}^{-1} \end{bmatrix} \begin{bmatrix} I & -\Sigma_{12}\Sigma_{22}^{-1} \\ \underline{0}' & I \end{bmatrix} \begin{bmatrix} \underline{x}_1-\underline{\mu}_1 \\ \underline{x}_2-\underline{\mu}_2 \end{bmatrix}$$

$$= \begin{bmatrix} \overbrace{\underline{x}_1-\underline{\mu}_1-\Sigma_{12}\Sigma_{22}^{-1}(\underline{x}_2-\underline{\mu}_2)}^{y_1} \\ \underbrace{\underline{x}_2-\underline{\mu}_2}_{y_2} \end{bmatrix}' \begin{bmatrix} (\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1} & \underline{0} \\ \underline{0}' & \Sigma_{22}^{-1} \end{bmatrix} \begin{bmatrix} \underline{x}_1-\underline{\mu}_1-\Sigma_{12}\Sigma_{22}^{-1}(\underline{x}_2-\underline{\mu}_2) \\ \underline{x}_2-\underline{\mu}_2 \end{bmatrix}$$

$$= y_1'(\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1}y_1 + y_2\Sigma_{22}^{-1}y_2$$

(c) The joint distribution of $\underline{x} = \begin{pmatrix} \overset{q\times 1}{\underline{x}_1} \\ \underset{(p-q)\times 1}{\underline{x}_2} \end{pmatrix}$ is :

$$f(\underline{x}) = \frac{1}{(2\pi)^{p/2}}\frac{1}{|\Sigma|^{1/2}}\exp\left(-\frac{1}{2}\left((\underline{x}-\underline{\mu})'\Sigma^{-1}(\underline{x}-\underline{\mu})\right)\right)$$

$$= \frac{1}{(2\pi)^{\frac{p-q}{2}}}\frac{1}{|\Sigma_{22}|^{1/2}}\exp\left(-\frac{1}{2}\left((\underline{x}_2-\underline{\mu}_2)'\Sigma_{22}^{-1}(\underline{x}_2-\underline{\mu}_2)\right)\right)$$

$$\times \frac{1}{(2\pi)^{q/2}}\frac{1}{|\Sigma_{11\cdot 2}|^{1/2}}\exp\left(-\frac{1}{2}y_1'\Sigma_{11\cdot 2}^{-1}y_1\right) \qquad (*)$$

where $\underset{\sim}{y}_1 = \underset{\sim}{x}_1 - \underset{\sim}{\mu}_1 - \Sigma_{12}\Sigma_{22}^{-1}(\underset{\sim}{x}_2 - \underset{\sim}{\mu}_2)$

$$\Sigma_{11\cdot 2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$$

Notice that the first term is the marginal distribution of $\underset{\sim}{x}_2$. By bayes theorem,

$$f(\underset{\sim}{x}) = f(\underset{\sim}{x}_2)\, f(\underset{\sim}{x} \mid x_2)$$

$$= f(\underset{\sim}{x}_2)\, f(\underset{\sim}{x}_1 \mid \underset{\sim}{x}_2) \qquad (**)$$

Compare $(*)$ and $(**)$

$$f(\underset{\sim}{x}_1 \mid \underset{\sim}{x}_2) = \frac{1}{(2\pi)^{q/2}} \frac{1}{|\Sigma_{11\cdot 2}|^{1/2}} \exp\left(-\frac{1}{2}\underset{\sim}{y}_1' \Sigma_{11\cdot 2}^{-1} \underset{\sim}{y}_1\right)$$

$$\Rightarrow \underset{\sim}{x}_1 \mid \underset{\sim}{x}_2 \sim N_q\left(\underset{\sim}{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\underset{\sim}{x}_2 - \underset{\sim}{\mu}_2),\ \Sigma_{11\cdot 2}\right)$$

Hotelling

4/19/2022

Package 'Hotelling'

Used to test the hypothesis of mean vectors for two groups:

$$\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$$

```
# install.packages('Hotelling') #uncomment this one to install package if you dont' have it
require(Hotelling)
data("container.df")
split.data = split(container.df[,-1], container.df$gp)
x = split.data[[1]]
y = split.data[[2]]
print(hotelling.test(x,y))
```

```
## Test stat:  2043
## Numerator df:  9
## Denominator df:  10
## P-value:  4.233e-09
```

Reject the null due to small p-value.

Confidence Region for mean vectors

```
#copied from E5-5r.tex
library(ellipse)
coltest=read.table("T5-2.DAT")
names(coltest)=c("SocialS","verbal","science")
n=nrow(coltest)
p=ncol(coltest)
c2=p*(n-1)*qf(.95,p,n-p)/(n-p)
xbar=colMeans(coltest)
S=cov(coltest)
 #calculate endpoints of axes
 eigv=eigen(S)
 a1=xbar-sqrt(eigv$value[1]*c2)%*%eigv$vector[,1]
 b1=xbar+sqrt(eigv$value[1]*c2)%*%eigv$vector[,1]
 a2=xbar-sqrt(eigv$value[2]*c2)%*%eigv$vector[,2]
 b2=xbar+sqrt(eigv$value[2]*c2)%*%eigv$vector[,2]
 a3=xbar-sqrt(eigv$value[3]*c2)%*%eigv$vector[,3]
 b3=xbar+sqrt(eigv$value[3]*c2)%*%eigv$vector[,3]
```

```
xl1=xbar[1]-sqrt(c2*S[1,1]/n)
xu1=xbar[1]+sqrt(c2*S[1,1]/n)
#for verbal
xl2=xbar[2]-sqrt(c2*S[2,2]/n)
xu2=xbar[2]+sqrt(c2*S[2,2]/n)
#for science
xl3=xbar[3]-sqrt(c2*S[3,3]/n)
xu3=xbar[3]+sqrt(c2*S[3,3]/n)
 #draw ellipse x1 x2
 xbar12=c(xbar[1],xbar[2])
 S12=matrix(c(S[1,1],S[2,1],S[1,2],S[2,2]),2,2)

eli = ellipse(S12, centre=xbar12,t=sqrt(c2/n), npoint=5000)

 plot(eli,
      cex=.3,
      bty="n",
      xlim=c(500,550),
      ylim=c(50,60),
      xlab="Social science",
      ylab="Verbal",type="l",
      lwd=2,col="blue")

segments(xl1,50-.5,xl1,xl2+1,lty=2)
segments(xu1,50-.5,xu1,xu2-1,lty=2)
segments(500-2,xl2,xl1+5,xl2,lty=2)
segments(500-2,xu2,xu1-5,xu2,lty=2)
```