# 多元正态的检验

王宁宁

需要安装的包：mvtnorm,ICP,ICSNP,biotools

## 多元随机向量的抽取

例 1：要从总体均值为（2，3），协方差阵为：

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

抽取 1000 个随机向量。

```
A=matrix(c(2,1,1,2),nrow=2,ncol=2);A
```

```
##      [,1] [,2]
## [1,]    2    1
## [2,]    1    2
```

```
library(mvtnorm)
set.seed(123)
X<-rmvnorm(1000, c(2, 3),A)
head(X)
```

```
##          [,1]     [,2]
## [1,] 1.150125 2.480423
## [2,] 4.155043 3.666843
## [3,] 2.804368 5.390145
## [4,] 2.166579 1.440601
## [5,] 0.898618 2.139809
## [6,] 3.803828 3.939560
```

```
cov(X)
```

```
##           [,1]      [,2]
## [1,] 1.8490873 0.8629169
## [2,] 0.8629169 2.0184171
```

```
cor(X)
```

```
##          [,1]     [,2]
## [1,] 1.000000 0.446668
## [2,] 0.446668 1.000000
```

```
summary(X)
```

```
##        V1                V2
##  Min.   :-1.999   Min.   :-0.762
```

```
##  1st Qu.: 1.124   1st Qu.: 2.083
##  Median : 2.144   Median : 3.061
##  Mean   : 2.060   Mean   : 3.041
##  3rd Qu.: 2.909   3rd Qu.: 4.042
##  Max.   : 6.858   Max.   : 8.328
```

# 多元正态性检验

例 2：检验上述例子是否来自正态总体。

```
library(ICS)
mvnorm.kur.test(X, method = "satt")
```

```
##
##  Multivariate Normality Test Based on Kurtosis
##
## data:  X
## W = 1.094, w1 = 1.5, df1 = 2.0, w2 = 2.0, df2 = 1.0, p-value =
## 0.8799
```

# 多元正态均值向量的检验

## 单个正态总体均值向量的检验

例 3：检验上面一个例子里的 X 是否来自均值向量为（2，3）的总体。

```
library(ICSNP)
HotellingsT2(X, mu=c(2,3))#Hotellings 检验#
```

```
##
##  Hotelling's one sample T2-test
##
## data:  X
## T.2 = 1.0386, df1 = 2, df2 = 998, p-value = 0.3543
## alternative hypothesis: true location is not equal to c(2,3)
```

```
HotellingsT2(X, mu=c(0,0))#Hotellings 检验#
```

```
##
##  Hotelling's one sample T2-test
##
## data:  X
## T.2 = 2483.7, df1 = 2, df2 = 998, p-value < 2.2e-16
## alternative hypothesis: true location is not equal to c(0,0)
```

例 4：教材 P39 3-4 题。

```
data3.3=read.csv("http://statstudy.github.io/data/data3_3.csv",head=T)#
读入数据#
head(data3.3)
```

```
##   X1 X2 X3 X4
## 1  5  6  9  8
## 2  8  5  3  6
## 3  9  6  7  9
## 4  9  2  2  8
## 5  9  4  3  7
## 6  9  5  3  7
```

```
HotellingsT2(data3.3, mu=c(7,5,4,8))#Hotellings 检验#
```

```
##
##  Hotelling's one sample T2-test
##
## data:  data3.3
## T.2 = 7.5968, df1 = 4, df2 = 16, p-value = 0.00125
## alternative hypothesis: true location is not equal to c(7,5,4,8)
```

## 两总体均值向量是否相等的检验

例 5： 检验一下下面随机生成的两个样本的均值向量是否相同。

```
X1<- rmvnorm(20, c(0, 0, 0, 0, 0))
X2<- rmvnorm(20, c(0,0, 4, 0, 0))
```

调用 Hotellings 检验

```
HotellingsT2(X1, X2)
```

```
##
##  Hotelling's two sample T2-test
##
## data:  X1 and X2
## T.2 = 30.443, df1 = 5, df2 = 34, p-value = 1.23e-11
## alternative hypothesis: true location difference is not equal to c(0,
0,0,0,0)
```

## 多个总体的均值向量的比较的检验

例 6：检验一下三个随机得到样本均值向量是否相同。

```
X1<- rmvnorm(20, c(0, 0, 0, 0, 0), diag(5))
X2<- rmvnorm(20, c(0,0, 4, 0, 0), diag(5))
X3<- rmvnorm(20, c(0,0, 0, 0, 0.05), diag(5))

X=rbind(cbind(1,X1),cbind(2,X2),cbind(3,X3));head(X)
```

```
##      [,1]       [,2]       [,3]       [,4]       [,5]       [,6]
## [1,]    1  2.2819670 -0.4636830 -0.3263536  0.8824932  1.2812861
## [2,]    1 -0.6586819  0.6645705 -0.5651575 -0.9621783  0.6233609
## [3,]    1  0.1064978  0.3893309 -0.5805035  1.7949780  0.6652880
## [4,]    1 -0.3744024  0.7027489 -1.2145144 -0.1377501  1.4033579
```

```
## [5,]    1 -0.1888393  0.9104904 -0.2219220 -2.2980264 -0.8802125
## [6,]    1  0.2227357  1.4465527 -0.5934021  0.2759790 -0.9648193
```

```
xx=X[,2:6]
head(xx)
```

```
##            [,1]       [,2]       [,3]       [,4]       [,5]
## [1,]  2.2819670 -0.4636830 -0.3263536  0.8824932  1.2812861
## [2,] -0.6586819  0.6645705 -0.5651575 -0.9621783  0.6233609
## [3,]  0.1064978  0.3893309 -0.5805035  1.7949780  0.6652880
## [4,] -0.3744024  0.7027489 -1.2145144 -0.1377501  1.4033579
## [5,] -0.1888393  0.9104904 -0.2219220 -2.2980264 -0.8802125
## [6,]  0.2227357  1.4465527 -0.5934021  0.2759790 -0.9648193
```

```
fac=factor(X[,1]);fac
```

```
##  [1] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2
 2 2 2
## [36] 2 2 2 2 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
## Levels: 1 2 3
```

```
re=manova(xx~factor(X[,1]))
summary(re)
```

```
##                Df Pillai approx F num Df den Df    Pr(>F)
## factor(X[, 1])  2  0.898   8.8007     10    108 1.988e-10 ***
## Residuals      57
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

## 多元总体的协方差阵的检验

例 7：检验一下上个例子中各个样本的协方差阵是否一致。

```
library(biotools)
```

```
## Loading required package: rpanel
## Loading required package: tcltk
## Package `rpanel', version 1.1-3: type help(rpanel) for summary infor
mation
## Loading required package: tkrplot
## Loading required package: MASS
```

```
## ---
## biotools version 2.2

##
```

```
boxM(X[,2:6],factor(X[,1]))
```

```
##
##  Box's M-test for Homogeneity of Covariance Matrices
##
```

```
## data:  X[, 2:6]
## Chi-Sq (approx.) = 41.993, df = 30, p-value = 0.07167
```